

OM of: US-09-810-936-304 to: GenEmbl.\* out\_format: pfs  
Date: Aug 26, 2002 4:26 AM  
About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:  
-MODEL=frimet.p2n.model -DEV=x1h  
-Q=/cgn2\_1/USPTO.spool/US09810936/runat.23082002.145849\_26171/app\_query.fasta\_1.450  
-DB=GenEmbl -OFMT=fastap -SUFFIX=p2n.rge -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOCPL=0.000 -LOCPEXT=0.000  
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -TGAPOP=10.000 -TGAPEXT=0.500  
-DELDP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blsuum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09810936.@CGN1.1.425 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPHY -WAIT -THREADS=1

Search information block:  
Query: US-09-810-936-304  
Query length: 384  
Database: GenEmbl.\*  
Database sequences: 1797656  
Database length: 187333701  
Search time (sec): 1851.490000

score_list:	stnd	orig	zscore	EScore	len	Documentation
gb_pat:AX106592	+ 2064.00	2864.74	3.3e-151	1155	1	AX106592 Sequence 373 from pat
gb_pat:AX140883	+ 2064.00	2864.74	3.3e-151	1155	1	AX140883 Sequence 373 from pat
gb_pat:AX200743	+ 2064.00	2864.74	3.3e-151	1155	1	AX200743 Sequence 373 from pat
gb_pat:AX267399	+ 2064.00	2864.74	3.3e-151	1155	1	AX267399 Sequence 373 from pat
gb_pat:AX282956	+ 2064.00	2864.74	3.3e-151	1155	1	AX282956 Sequence 301 from pat
gb_pat:AX316964	+ 2064.00	2861.82	4.8e-151	1590	1	AX316964 Sequence 301 from pat
gb_pat:AX316986	+ 2054.00	2850.80	2.0e-150	1155	1	AX316986 Sequence 323 from pat
gb_pat:AX316991	+ 2024.00	2804.04	8.0e-148	2000	1	AX316991 Sequence 374 from pat
gb_pat:AX106593	+ 2024.00	2804.04	8.0e-148	2000	1	AX106593 Sequence 374 from pat
gb_pat:AX140884	+ 2024.00	2804.04	8.0e-148	2000	1	AX140884 Sequence 374 from pat
gb_pat:AX200744	+ 2024.00	2804.04	8.0e-148	2000	1	AX200744 Sequence 374 from pat
gb_pat:AX267400	+ 2024.00	2804.04	8.0e-148	2000	1	AX267400 Sequence 374 from pat
gb_pat:AX282957	+ 2024.00	2804.04	8.0e-148	2000	1	AX282957 Sequence 6 from pat
gb_pat:AX316965	+ 2024.00	2804.04	8.0e-148	2000	1	AX316965 Sequence 302 from pat
gb_pat:AX106594	+ 2024.00	2803.85	8.2e-148	2040	1	AX106594 Sequence 375 from pat
gb_pat:AX140885	+ 2024.00	2803.85	8.2e-148	2040	1	AX140885 Sequence 375 from pat
gb_pat:AX200745	+ 2024.00	2803.85	8.2e-148	2040	1	AX200745 Sequence 375 from pat
gb_pat:AX267401	+ 2024.00	2803.85	8.2e-148	2040	1	AX267401 Sequence 375 from pat
gb_pat:AX282958	+ 2024.00	2803.85	8.2e-148	2040	1	AX282958 Sequence 7 from pat
gb_pat:AX316966	+ 1920.50	2660.74	7.7e-140	1853	1	AX316966 Sequence 303 from pat
gb_pat:AX140879	+ 1920.50	2660.74	7.7e-140	1853	1	AX140879 Sequence 369 from pat
gb_pat:AX200739	+ 1920.50	2660.74	7.7e-140	1853	1	AX200739 Sequence 369 from pat
gb_pat:AX267395	+ 1920.50	2660.74	7.7e-140	1853	1	AX267395 Sequence 369 from pat
gb_pat:AX316958	+ 1876.00	2603.17	1.2e-136	1155	1	AX316958 Sequence 329 from pat
gb_pat:AX316988	+ 1857.00	2576.73	3.7e-135	1155	1	AX316988 Sequence 325 from pat
gb_pat:AX316993	+ 1857.00	2576.73	3.7e-135	1155	1	AX316993 Sequence 330 from pat
gb_pat:AX141040	+ 1495.00	2068.74	7.2e-107	1852	1	AX141040 Sequence 530 from pat
gb_pat:AX200900	+ 1495.00	2068.74	7.2e-107	1852	1	AX200900 Sequence 530 from pat
gb_pat:AX267556	+ 1495.00	2068.74	7.2e-107	1852	1	AX267556 Sequence 530 from pat
gb_pat:AX282954	+ 1495.00	2068.74	7.2e-107	1852	1	AX282954 Sequence 313 from pat
gb_pat:AX316976	+ 1469.00	2034.43	5.9e-105	1512	1	AX316976 Sequence 368 from pat
gb_pat:AX106587	+ 1469.00	2034.43	5.9e-105	1512	1	AX106587 Sequence 368 from pat
gb_pat:AX140878	+ 1469.00	2034.43	5.9e-105	1512	1	AX140878 Sequence 368 from pat
gb_pat:AX200738	+ 1469.00	2034.43	5.9e-105	1512	1	AX200738 Sequence 368 from pat
gb_pat:AX267394	+ 1469.00	2034.43	5.9e-105	1512	1	AX267394 Sequence 368 from pat
gb_pat:AX316957	+ 1468.00	2031.18	8.9e-105	1851	1	AX316957 Sequence 291 from pat
gb_pat:AX148112	+ 1468.00	2031.18	8.9e-105	1851	1	AX148112 Sequence 291 from pat
gb_pat:AX106585	+ 1468.00	2031.18	8.9e-105	1851	1	AX106585 Sequence 366 from pat

gb\_pat:AX140876 - 1468.00 2031.18 8.9e-105 1851 1 AX140876 Sequence 366 from  
gb\_pat:AX200736 - 1468.00 2031.18 8.9e-105 1851 1 AX200736 Sequence 366 from  
gb\_pat:AX267392 - 1468.00 2031.18 8.9e-105 1851 1 AX267392 Sequence 366 from  
gb\_pat:AX282952 - 1468.00 2031.18 8.9e-105 1851 1 AX282952 Sequence 1 from Pa  
gb\_pat:AX316954 - 1468.00 2031.18 8.9e-105 1851 1 AX316954 Sequence 291 from

seq\_name: gb\_pat:AX106592  
seq\_documentation\_block: 1155 bp DNA linear PAT 30-APR-2001  
LOCUS AX106592  
DEFINITION Sequence 373 from Patent WO0125272.  
ACCESSION AX106592  
VERSION AX106592.1 GI:13922263  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.  
TITLE Compositions and methods for therapy and diagnosis of prostate  
cancer  
JOURNAL Patent: WO 0125272-A 373 12-APR-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source location/Qualifiers  
1..1155  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 346 a 253 c 297 g 259 t  
ORIGIN

alignment\_scores:  
Quality: 2064.00 Length: 384  
Ratio: 5.375 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-810-936-304 x AX106592 ..

Align seg 1/1 to: AX106592 from: 1 to: 1155

1 MetValValGluValAspSerMetProAlaMetSerValLysLysPr 17  
|||||  
1 ATGctGctGAGctGctGctATCCATGcGcGctGctTCTTGtGAAGAGcC 50  
|||||  
17 ophegLyLeuArgSerLysMetGlyLysTrpCysArgCysPheProc 34  
|||||  
51 ATTGTGCTCAGAGAGAGATGGCAGTGGCTGGCTGCTTCCCT 100  
|||||  
34 yscsArgLysSerGlyLysSerAsnValGlyThrSerGlyAspHisasp 50  
|||||  
101 GCTGcAGGAGAGcGcAGcAGcAGcAGcAGcAGcAGcAGcAGcAGcAGc 150  
|||||  
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTrpCysArgH 67  
|||||  
151 GACTGTGCTATGAGACACTGAGAGCAAGATGGCAGTGGCTGCGCGCA 200  
|||||  
67 scysPheProcCysArgGlySerGlyLysSerAsnValGlyLysasp 84  
|||||  
201 CTGCTTCCCTCTCTCAGAGGAGTGGcAGcAGcAGcAGcAGcAGcAGc 250  
|||||  
84 LysPheLysAspSerAlaMetLysThrLeuArgSerLysMetGlyLys 100  
|||||  
251 GAGACcAGcAGcAGcAGcAGcAGcAGcAGcAGcAGcAGcAGcAGcAG 300  
|||||  
101 TrpCysCysHisCysPheProcCysArgGlySerGlyLysSerLysVa 117  
|||||  
301 TGGTGTGCTGCTGCTCTCCCTGCTGcAGGAGcAGcAGcAGcAGcAG 350  
|||||  
117 LcGlyAlaTrpGlyLysTrpAspAspSerAlaPheMetLysLysProGly 134  
|||||  
351 GGGCGCTGGGAGcAGcAGcAGcAGcAGcAGcAGcAGcAGcAGcAGcAG 400

134 isValArgGlyGluAspLeuAspLysLeuHisArgAlaAlaTrpTrpGly 150  
|||||  
401 ACGTCCGCGGAGAGATCTGGACAAGCTCCACAGAGCTGCTGGTGGGGT 450  
151 LysValProArgLysAspLeuIleValMetLeuArgAspThrAspValAs 167  
|||||  
451 AAAGTCCCGAGAAAGATCTCATGCTCAGGAGGACTGACGTGAA 500  
167 nLysLysAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsn 184  
|||||  
501 CAAAGAGACAAAGCAAAAGAGACTGCTCATCTGAGCCCTGCTCCATG 550  
184 LysAsnSerGluValValLysLeuLeuAsnAspArgArgCysGlnLeuAsn 200  
|||||  
551 GGAATTCAGAAAGTAGTAATAACTCCCTGCTGACAGACGATGCAACTTAAT 600  
201 ValLeuAspAsnLysLysArgThrAlaLeuIleLysAlaValGlnCysGln 217  
|||||  
601 GTCTTTGACAAACAAAAGAGAGACGCTCTGATAAAGCCGTACAAATGCCA 650  
217 nGluAspGluCysAlaLeuMetLeuLeuGlnHisGlyThrAspProAsnI 234  
|||||  
651 GGAAGATGATGTGGCTTAATGTTGCTGGAACATGACACTGATCCAAATA 700  
234 LeProAspGluTrpGlyLysAsnThrThrLeuHisTyrAlaIleGlyAsnGlu 250  
|||||  
701 TTCAGATGATGATGAAATACCACTGCTGCACTACGCTATCATATATATA 750  
251 AspLysLeuMetAlaLysAlaLeuLeuLeuTrpGlyAlaAspIleGluSe 267  
|||||  
751 GATTAATTAATGGCCAAAGCACTGCTTATATGTTGCTGATATGCAATG 800  
267 rLysAsnLysHisGlyLeuThrProLeuLeuLeuGlyValHisGlnGln 284  
|||||  
801 AAAAAACAAGCATGGCTTCACACCACTGTTACTGCTGACATGAGCAAA 850  
284 yGlnGlnValValLysPheLeuIleLysLysLysAlaAsnLeuAsnAla 300  
|||||  
851 AACAGCAATGCTCGAATTTTATATCAAGAAAAAGCGAATTTAATATGCA 900  
301 LeuAspArgTrpGlyLysArgThrAlaLeuIleLeuAlaValCysGlySe 317  
|||||  
901 CTGGATAGATATGGAAGAGACTGCTCATATCTGCTGTTGTTGGATC 950  
317 rAlaSerIleValSerLeuLeuLeuGlnGlnAsnIleAspValSerSerG 334  
|||||  
951 AGCAAGATATGTCAGCCTTCTACTTGAGCAAAATATTGATGATCTTCTC 1000  
334 InAspLeuSerGlyGlnThrAlaArgGluTrpAlaValSerSerHisHis 350  
|||||  
1001 AAGATCTATCTGACAGACGGCCAGAGAGATATGCTTCTTAGTCATCAT 1050  
351 HisValIleCysGlnLeuLeuSerAspTrpLysGlnLysGlnMetLeu 367  
|||||  
1051 CATGTATATTTGCCAGTTACTTCTGACTACAAAAGAAAAACAGATGCTAAA 1100  
367 sIleSerSerGluAsnSerAsnProGluAsnValSerArgThrArgAsnL 384  
|||||  
1101 AATCTCTTCTGAAAACAGCAATCCAGAAATGTCTCAAGAACCAAGAAATA 1150  
384 ys 384  
||  
1151 AA 1152

seq\_name: gb\_pat:AX140883  
seq\_documentation\_block: 1155 bp DNA linear PAT 31-MAY-2001  
LOCUS AX140883  
DEFINITION Sequence 373 from Patent WO0134802.  
ACCESSION AX140883  
VERSION AX140883.1 GI:14280986  
KEYWORDS

SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
1 (bases 1 to 1155)  
AUTHORS  
Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,  
Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,  
Skelly,X.A. and Wang,A.  
TITLE  
Compositions and methods for the therapy and diagnosis of prostate  
cancer  
JOURNAL  
Patent: WO 0134802-A 373 17-MAY-2001;  
CORIXA CORPORATION (US)  
FEATURES  
Location/Qualifiers  
source 1..1155  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 346 a 253 c 297 g 259 t  
ORIGIN  
alignment\_scores:  
Quality: 2064.00 Length: 384  
Ratio: 5.375 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-810-936-304 x AX140883 ..  
Align seg 1/1 to: AX140883 from: 1 to: 1155  
1 MetValValGluValAspSerMetProAlaIleSerSerValLysLysPr 17  
|||||  
1 ATGGTGGTTGAGGTTGATTCATCCATGCGCGCTGCTCTGTGTAAGAACCC 50  
17 oPheGlyLeuArgSerLysMetGlyLysTrpCysArgCysPheProC 34  
|||||  
51 ATTGGTCTCAGAGCAAGATGGGCAAGTGTGCTGCTGCCCTCCCT 100  
34 yScysArgGluSerGlyLysSerAsnValGlyThrSerGlyAspHisAsp 50  
|||||  
101 GCTCAGAGGAGAGCGGCAGAGCAAGCTGGCACTTCTGGAGACCAAGAC 150  
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTrpCysArgH 67  
|||||  
151 GACTCTGCTATGAAACACTCAGAGCAAGATGGGCAAGTGTCCGCCA 200  
67 sCysPheProCysCysArgGlySerGlyLysSerAsnValGlyAlaSerG 34  
|||||  
201 CTGCTTCCCTGTCGACAGGGGAGCTGCAAGAGCAAGCTGGCGCTCTCTG 250  
84 LysAspHisAspAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100  
|||||  
251 GAGACACAGAGCACTGCTGTATGAAGACACTCAGAAACAGATGGGCAAG 300  
101 TrpCysCysHisCysPheProCysCysArgGlySerGlyLysSerLysVa 117  
|||||  
301 TGGTCTGCTCCACTGCTTCCCTGCTGACGGGAGCGGCAAGACCAAGGT 350  
117 lGlyAlaTrpGlyAspTrpAspAspSerAlaPheMetGluProArgTrpH 134  
|||||  
351 GGGCGCTTGGGGAGACTACGATGACAGTGCCTTCATGAGGCCACGATACC 400  
134 isValArgGlyGluAspLeuAspLysLeuHisArgAlaAlaTrpTrpGly 150  
|||||  
401 ACGTCCGCGGAGAGATCTGGACAAGCTCCACAGAGCTGCTGGTGGGGT 450  
151 LysValProArgLysAspLeuIleValMetLeuArgAspThrAspValAs 167  
|||||  
451 AAAGTCCCGAGAAAGATCTCATGCTCAGAGGACACTGACGTGAA 500  
167 nLysLysAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsn 184  
|||||  
501 CAAAGAGACAAAGCAAAAGAGACTGCTCTACATCTGCGCTGCAATG 550

184 lysnserctluvallyallysleuleuleuaspragdygsglnleuasn 200  
|||||  
551 ggatttcagaaattag7aaataactctcttgagacagcagttgcaacttaatt 600  
|||||  
201 valleuaspsanlyslysargtthralaleu1lelvasalavalgincysgi 217  
|||||  
601 gtctcttgacaaacaaaaagacagacctgcgtaaaagcccgacaaagcca 650  
|||||  
217 nguaaspglucysalaleuemetleuenu1un1sglythrasproasn1 234  
|||||  
651 ggaaagtgaatgctgcttaattgcttggaacatgacgcactgaatccaaata 700  
|||||  
234 letrroaspolutrglyasnthrthlenn1styyala1leltyaaanglu 250  
|||||  
701 ttccaatatgagatgaaatpccacacctgactacgctatctaatatga 750  
|||||  
251 asplysleumetalalysalaleu1leu1trgilyalaaspi1leguise 267  
|||||  
751 gataaattatattggccaaagcacgcctcttattatgctgcgtatnccatc 800  
|||||  
267 rlysaenlyshtsglyleuthrproleu1leuenu1yala1sglglrl 284  
|||||  
801 aaaaaaacagcatggcctcacacacactgttacttgctgacatagcama 850  
|||||  
284 ysglndlnvallyalysphelenu1lelysllysalasnlleuasn1a 300  
|||||  
851 aacagcaagtcgtgaattttatnatacaaaaaaaccaattttaaatga 900  
|||||  
301 leuaspargtyrglyargthralaleu1leleuala1alycysgylse 317  
|||||  
901 ctggaatagatagcagagacacgcctctcattacatgctgattgttgatc 950  
|||||  
317 talaser1levalserleu1leuenglun1asn1leaspralserse1s 334  
|||||  
951 agcaatgatagtcacaccttctacttgagcaaaatattgatgatcttctc 1000  
|||||  
334 lnspleuserglvg1trthralargglutyr1alavalserse1sh1s 350  
|||||  
1001 aagatctatctgacacagcccgaaagatn1gctgttctagtcacatc 1050  
|||||  
351 hisval1lecyglnleu1leuseraspyrtyrlysgulysglnmetleu1y 367  
|||||  
1051 catgtaattttggcagcttaccttgcactcaaaagaaaaacgatgctaaa 1100  
|||||  
367 s1ieserSergluaenSerasnprogluaenvalser1trgthrragsnl 384  
|||||  
1101 aatctcttctgaaaaacacacatccagaaatgcttcaagaacacagaaata 1150  
|||||  
384 ys 384  
||  
1151 aa 1152

seq\_name: gb\_pat:AX200743

seq\_documentation\_block: 1155 bp DNA linear PAT 29-AUG-2001

LOCUS AX200743  
DEFINITION Sequence 373 from Patent WO0151633.  
VERSION AX200743  
ACCESSION AX200743  
KEYWORDS AX200743.1 GI:15390636

ORGANISM human.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,  
Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Ketter,M.W.,  
Stolk,J.A., Skeiky,Y.A., Wang,A. and Weaghter,M.J.  
Compositions and methods for the therapy and diagnosis of prostate  
cancer  
Patent: WO 0151633-A 373 19-JUL-2001;  
CORIXA CORPORATION (US)

FEATURES	Location/Qualifiers
source	1..1155
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
BASE COUNT	346 a 253 c 297 g 259 t
ORIGIN	
alignment_scores:	
Quality:	2064.00
Ratio:	5.375
Percent Similarity:	100.000
Percent Identity:	100.000
alignment_block:	
US-09-810-936-304 x AX200743	..
Align seg 1/1 to: AX200743 from: 1 to: 1155	
1	MetValaIGluValaAspSerMetProAlaIaSerValIysIysPr 17
1	ATNGTGGTGGAGGTTGATTCCATGCGGGCTCCCTCTCGGAAGAAC 50
17	obheGlyLeuArGserIysMetGlyIysTrpCysArgCysPheProC 34
51	ATTGGCTCGAGACCAAGATGGGAAGAGTGGTGGCCGTGGCTCCCT 100
34	ysCysArgIuSerGlyIysSerAsnValGlyThrSerGlyAspPHisAsp 50
101	GGTGAGAGGAGACGGGCAAGAACCGTGGCACTCTGGAGACCCAGAC 150
51	AspSerAlaMetIysThrLeuArgSerIysMetGlyIysTrpCysArgH 67
151	GACTCTGCATTAACAACACTCAGAGAGCAAGATGGGCAAGTGTCCCCA 200
67	sCysPheProCysCysArgIysSerGlyIysSerAsnValGlyAlaSerG 84
201	GTGCTTCCCTTCCTCAGAGGAGTGGCAAGACAGACGTGGCCCTTCTG 250
84	IysPHisAspSerAlaMetIysThrLeuArgAsnIysMetGlyIys 100
251	GAGACACACAGCACTCTGATGAAGACACTCAGAGAACAGATGGGCAG 300
101	TrpCysCysHHisCysPheProCysCysArgIysSerGlyIysSerIysVa 117
301	TGGTGGTGGCACTGCTTCCCTCGCTGGAGGGGAGAGCGCAAGACCAAGT 350
117	IcGlyAlaTrpGlyAspArgIysAspAspSerAlaPheMetGluProArgIyrH 134
351	GGGCGCTTGGGAGACTCGATGACAGAGCCTTCAATGGAGCCACGGTACC 400
134	IaValArgIyGluAspLeuAspIysLeuHISArgAlaIaIaTrpTrpGly 150
401	ACGTCCGTGGAGAAATGATCTGGACAAGCTCCACAAAGTGGCTGTGGGGT 450
151	LysValProArgIysAspLeuIleValMetLeuArgAspTrpAspValAs 167
451	AAAGTCCCAAGAAAGATCTCATCTCATGCTCAAGGAGCACTGACGGAA 500
167	nIysIysAspIysGlnIysArgTrhAlaLeuHISLeuAlaSerAlaansG 184
501	CAGAAGAGCAAGCAAAAGAGAGACAGCTCTCAATCTGGCCCTTGGCAATG 550
184	IysAsnSerIuValValIysLeuLeuLeuAspArgArgCysGlnLeuAsn 200
551	GGAAATCACAAGATGAGAAACTCTCGTGGACAGACGATGCACCTTAAT 600
201	ValIleuAspArgIysArgTrhAlaLeuIleLysAlaValAlaGlnCysG 217
601	GTCTTGGACACAAAGAGGACAGCTGTGATAAAGCCGATACATGCCA 650
217	nGluAspArgIuysAlaLeuMetLeuGlnHISGlyThrAspProAsnI 234
651	GGAAAGTGAATGTGGCTTAATCTCTCTGGACATGGCCATCATCAATA 700

```

234 leProaspGluTyrGlyAsnThrThrLeuHisTyrAlaIleTyrAsnGlu 250
|||||
701 TTCAGATGAGTATGGAATACCACTGCTGCACCTACCGCATATATAATGAA 750
251 AspLysLeuMetAlaLysAlaLeuLeuTyrGlyAlaAspIleGluSe 267
|||||
751 GATTAATTAATGAGCAAGACACTGCTTATATAGTGCTGATATACGAAATC 800
267 rLysAsnLysHisGlyLeuThrProLeuLeuLeuGlyValHisGluInL 284
|||||
801 AAAAACAACAGCATGGCTCACAACACTGTTACTTGCTGATACATGAGCAA 850
284 ysgInGInValAlaLysPheLeuIleLysLysLysAlaAsnLeuAsnAla 300
|||||
851 AACGGCAGTCGTGAAATTTTAAATCAGAAAGCAATTAATTAATGCA 900
301 LeuAspArgTyrGlyArgThrAlaLeuIleLeuAlaValCysCysGlySe 317
|||||
901 CTGGATGATATGGAAGACTGCTCTCATCTGCTGATGCTGTGATC 950
317 rAlaSerIleValSerLeuLeuLeuGluInLAsnIleAspValSerSerg 334
|||||
951 AGCAGATATGTCAGCCTCTACTTGAGCAAAATATGATATCTCTC 1000
334 lnaAspLeuSergGlyInThrAlaATGGLuTyrAlaValSerSerHis 350
|||||
1001 AAGATCTATCTGGACAGACGGCCAGAGATATGCTTTCTATGTCATCAT 1050
351 HisValIleCysGlnLeuLeuSerAspTyrLysGluLysGlnMetLeu 367
|||||
1051 CATCTAATTTGCCAGTTACTTCTGACTACAAAGAAAAAGATGCTAAA 1100
367 sIleSerSergLysAsnSergAsnProGluAsnValSerArgThrArgAsn 384
|||||
1101 AATCTCTTCGAAACAGCAATCCAGAAATATCTCAAGAAACCGAAATA 1150
384 ys 384
||
1151 AA 1152
seq_name: gb_pat:AX267399
seq_documentation_block: 1155 bp DNA linear PAT 26-OCT-2001
LOCUS AX267399
DEFINITION Sequence 373 from Patent WO0173032.
ACCESSION AX267399
VERSION AX267399.1 GI:16516169
KEYWORDS
SOURCE
ORGANISM human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (sites)
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,X.,
Kados,M.D., Fanger,G.R., Rether,M.W., Stolk,J.A., Day,C.H.,
Vedrick,T.S., Carter,D., Li,S.X., Wang,A., Skelky,Y.A., Hepler,W.T.
and Henderson,R.A.
TITLE Compositions and methods for the therapy and diagnosis of prostate
cancer
JOURNAL Patent: WO 0173032-A 373 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source 1..1155
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 346 a 253 c 297 g 259 t
ORIGIN
alignment_scores:
Quality: 2064.00 Length: 384
Ratio: 5.375 Gaps: 0

```

```

Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-810-936-304 x AX267399
Align seg 1/1 to: AX267399 from: 1 to: 1155
1 MetValValGluValAspSerMetProAlaAlaSerValLysLysPr 17
1 ATGCTGCTTGAGGTTGATTCATGCGCGCTGCTCTCTCTGTAAGAGCC 50
17 opheGlyLeuArgSerLysMetGlyLysTrpCysArgCysPheProc 34
51 ATTGGTCTCAGAGCAAGATGGCAAGTGTGCTGCGCTTGTCTCCCT 100
34 yscYsArgLysSerGlyLysSerAsnValGlyThrSergGlyAspHisAp 50
101 GCTGCAGGAGAGCGGCAAGACACGTGGGCACTTCTGAGACACGAC 150
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTrpCysArgH 67
151 GACTCTGCTATGAGACACTCAGAGCAAGATGGGCAAGTGGCGGCA 200
67 sCysPheProcCysArgGlySergLysSerAsnValGlyAlaSerg 84
201 CTGCTTCCCTCTGCGAGGGGAGTGGCAAGACACGTGGGCGCTTCG 250
84 lYsAspHisAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100
251 GAACACACAGACACCTCTGTATGAAGACATCAGAAACAAGATGGGCAAG 300
101 TrpCysCysHisCysPheProcCysArgGlySergLysSerLysVal 117
301 TGATGTGTCACACTGCTCCCTGCTGTCAGGGGAGCGGCAAGAGCAAG 350
117 lGlyAlaTrpGlyAspTyrAspAspSerAlaPheMetGlnProArgTyrH 134
351 GGGCGCTGGGAGACTACGATGACAGTGCCTCATGAGCCCAAGGTACC 400
134 lValArgGlyLysLysLeuAspLysLysLeuHisArgAlaAlaTrpTrpGly 150
401 ACGTCCGAGAGAGATCGGACAGGTCCACAGAGCTGCTGATGGGGT 450
151 lYsValProArgLysAspLeuIleValMetLeuArgAspHisArgValAs 167
451 AAAGTCCCAAGAAAGATCTCATCTGCTCAGGGACACTGACGTGAA 500
167 nLysLysAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsn 184
501 CAAGAAAGACAAAGAAAGAGACTGCTCTACATCTGCGCTGCGCAATG 550
184 lYsAsnSergLysValValLysLeuLeuAspArgArgCysGlnLeuAsn 200
551 GGAATTCAGAGTAGTAATAACTCTGCTGAGACAGACATGTCAACTAAT 600
201 ValLeuAspAsnLysLysArgThrAlaLeuIleLysAlaValGlnCysG 217
601 GTCCTTGACAAACAAAAGAGACACTGTGAATAAGGCCCTGACAATGCCA 650
217 nGluAspGlyLysAlaLeuMetLeuGlnHisGlyThrAspProAsnI 234
651 GGAAGATGAATGGCGTTAATGTTCTGGAACATGGCAGATGACCAATA 700
234 leProaspGluTyrGlyAsnThrThrLeuHisTyrAlaIleTyrAsnGlu 250
701 TTCAGATGAGTATGGAATACCACTGCTGCACCTACCGCATATATAATGAA 750
251 AspLysLeuMetAlaLysAlaLeuLeuTyrGlyAlaAspIleGluSe 267
751 GATTAATTAATGAGCAAGACACTGCTTATATAGTGCTGATATACGAAATC 800
267 rLysAsnLysHisGlyLeuThrProLeuLeuLeuGlyValHisGluInL 284

```

```
801 AAAAAACAGCATGGCCTCACACCCTGTTACTTGTTGATGACCAAA 850
284 ysginglnvalValIysPheleuileYsAlaAsnLeuAna1a 300
851 AACGACAGTCGTGAAATTTTAAATCAAGAAAAAGCAATTAATGCA 900
301 LeuAspArgTyrGlyArgThrAlaLeuileLeuAlaValCysCysGlySe 317
901 CTGGATAGATATGAAAGACCTGCTCATACTTGCTGTATGTTGTGATC 950
317 fAlaSerIleValSerIleuLeuGluGlnAsnIleAspAlaSerSerg 334
951 AGCAAGATATGTCAGCCTTCTACTTGAGCAAAATTCATGATCATCTTC 1000
334 lnaAspLeuSerGlyGlnThrAlaArgLutYrAlaValSerSerHis 350
1001 AAGATCATCTGACAGAGCGCCAGAGATGCTGTTCTAGTCATCAT 1050
351 HisValIleCysGlnLeuLeuSerAspTyrLysGluLysGlnMetLeu 367
1051 CATGTAAATTTGCCACTTACTTCTGACTACAAAGAAAAACGATGCTAA 1100
367 sIleSerSerGluAsnSerAsnProGluAsnValSerArgThrArgAsn 384
1101 AATCTCTCTGAAAAACAGCAATCCAGAAATGTCTCAGAACGAAKHA 1150
384 ys 384
1151 AA 1152
seq_name: gb_pat:AX282956
seq_documentation_block: 1155 bp DNA linear PAT 02-NOV-2001
LOCUS AX282956 1155 bp
DEFINITION Sequence 5 from Patent WO01/75171.
ACCESSION AX282956
VERSION AX282956.1 GI:16609896
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (sites)
AUTHORS
Houghton, R.L., Dillon, D.C., Molesh, D.A., Xu, J., Zehentner, B. and
Pearling, D.H.
TITLE
Methods, compositions and kits for the detection and monitoring of
breast cancer
JOURNAL
Patent: WO 01/75171-A 5 11-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source
1. 1155
Location/Qualifiers
BASE COUNT 346 a 253 c 297 g 259 t
ORIGIN
Alignment_scores:
Quality: 2064.00 Length: 384
Ratio: 5.375 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-810-936-304 x AX282956 ..
Align seg 1/1 to: AX282956 from: 1 to: 1155
1 MetAlaValGluValAspSerMetProAlaAlaSerSerValLysLysPr 17
|||||
1 ATGCTGCTTGAAGTGTGATTCATCCGCGCTCTCTCTGCAAGAACGCC 50
|||||
17 opbGlyLeuArgSerLysMetGlyLysTyrPyrCysArgGlySphProC 34
|||||
```

```
51 AATTGCTCTCAGACCAAGATGGCAAGTGGTGTGCTGCCGTTGCTCCCT 100
34 yScysArgGluSerGlyLysSerAsnValGlyThrSerGlyAspHisAsp 50
101 GCTCAGAGGAGAGCGCAAGAGCAACCTGGGCTTCTGGAGACCAAGAC 150
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTyrPyrCysArg 67
151 GACTCTGCTATGAGACACTCAGCAGCAAGATGGCCAAAGTGTGCCGCC 200
67 sCysPheProCysCysArgGlySerGlyLysSerAsnValGlyAlaSerG 84
201 CTGCTTCCCTGCTGCGAGGGGAGTGGCAAGACCAAGCTGGCGCTTC 250
84 lYAspHisAspAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100
251 GAGACCAAGACGACTCTGCTATGAAAGCACTCAGAACCAAGATGGCGA 300
101 TyrPyrCysHisCysPheProCysCysArgGlySerGlyLysSerLysVa 117
301 TGCTGCTGCCACTGCTCCCTGCTGCGAGGGGAGCGCAAGACCAAGGT 350
117 lGlyAlaThrGlyAspTyrAspAspSerAlaPheMetGluProArgTyrH 134
351 GAGCGCTTGGGAGACTACGATGACAGTGTCTTCATGAGAGCCCAAGTAC 400
134 lSValArgGlyGlyLysPheuAspLysLeuHisArgAlaAlaTrpTyrGly 150
401 ACGTCCGTGGAGAGATCTGACAAAGCTCCACAGAGCTGCTGGTGGGG 450
151 LysValProArgLysAspLeuileValMetLeuArgAspThrAspValAs 167
451 AAATCTCCCGAAAGATCTCATCTGATGCTTCAGGACACTGACGTGGA 500
167 nLysLysAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsn 184
501 CAAGAAGAGCAACAAAGAGAGACTGCTACACTGCGCTCTCCCAATG 550
184 lYAsnSerGluValValLysLeuLeuLeuAspArgArgCysGlnLeuAsn 200
551 GGAATTCAGAGAGTAAACTCTGCTGAGCAGAGATGTCACACTTAAT 600
201 ValLeuAspAsnLysLysArgThrAlaLeuileLysAlaValGlnCysG 217
601 GTCTTGACAAACAAAAGAGAGAGCTGTGATTAAGGCCGACCAATCC 650
217 nGluAspGluCysAlaLeuMetLeuGluHisGlyThrAspProAsn 234
651 GGAAGATGAATGTGCCTTAATGTTGCGAACAATGGCACTGATCCAAAT 700
234 lProAspGluTyrGlyAsnThrThrLeuHisTyrAlaAlaTyrAsnGlu 250
701 TTCAGATGAGTATGGAATATCCACTCTGCACACTACCTATATATGA 750
251 AspLysLeuMetAlaLysAlaLeuLeuLeuTyrGlyAlaAspIleGluSe 267
751 GATTAATTAATGSCCAAGACACTCTCTATATGTCGCGATATCGAATC 800
267 lYsAsnLysHisGlyLeuThrProLeuLeuGluGlyValHisGluLnt 284
801 AAAAAACAGCATGGCCTCACACCCTGTTACTTGTTGATGACCAAA 850
284 ysginglnvalValIysPheleuileYsAlaAsnLeuAna1a 300
851 AACGACAGTCGTGAAATTTTAAATCAAGAAAAAGCAATTAATGCA 900
301 LeuAspArgTyrGlyArgThrAlaLeuileLeuAlaValCysCysGlySe 317
901 CTGGATAGATATGAAAGACCTGCTCATACTTGCTGTATGTTGTGATC 950
317 fAlaSerIleValSerIleuLeuGluGlnAsnIleAspAlaSerSerg 334
951 AGCAAGATATGTCAGCCTTCTACTTGAGCAAAATTCATGATCATCTTC 1000
```

```
334 1nAspLeuSerGlyGlnThrAlaArgLutryAlaValSerSerHis 350
|||||
1001 AAGATCTATGAGACAGCGCCAGAGATGCTGTTCTTCAAGCATCAT 1050
|||||
351 HisValIleCysGlnLeuLeuSerAspTyrLysGlnLysGlnMetLeu 367
|||||
1051 CATGTAATTTGCCAGTTACTTCTTGACTCAAGAAAGAAAGATGCTAAA 1100
|||||
367 sIleSerSerGluAsnSerAsnProGluAsnValSerArgThrArgAsn 384
|||||
1101 AATCTCTCTGAAACAGCAATCCAGAAATGCTCAAGAACAGAAATG 1150
|||||
384 ys 384
|||
1151 AA 1152

seq_name: gb_pat:AX316964
seq_documentation_block: 1155 bp DNA linear PAT 14-DEC-2001
LOCUS AX316964
DEFINITION Sequence 301 from Patent WO0190152.
ACCESSION AX316964
VERSION AX316964.1 GI:17900043
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (sites)
Fudakis, J.N., Reed, S.G., Smith, J.M., Misher, L.E., Dillon, D.C.,
Retter, M.W., Wang, A., Skelky, J.A., Harlocker, S.L. and Day, C.H.,
Compositions and methods for the therapy and diagnosis of breast
cancer.
Patent: WO 0190152-A 301 29-NOV-2001;
JOURNAL CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1. 1155
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 346 a 253 c 297 g 259 t
ORIGIN
Alignment_scores:
Quality: 2064.00 Length: 384
Ratio: 5.375 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
us-09-810-936-304 x AX316964 ..
Align seg 1/1 to: AX316964 from: 1 to: 1155
1 MetValValGluValAspSerMetProAlaIleAspSerValLysLysP 17
|||||
1 ATGTTGGTTGAGGTGATTCATCCCGCTGCTCTCTCGTAGAAGAGCC 50
|||||
17 oPhgGlyLeuArgSerLysMetGlyLysTyrPyrCysArgCysPheProC 34
|||||
51 ATTTGGTCTCAGAGCAGAGATGGCCAGTGGTCTGCTCCCTTCTCCCT 100
|||||
34 yscsArgGluSerGlyLysSerAsnValGlyThrSerGlyAspHisAsp 50
|||||
101 GCTCGAGGAGAGCGGCAAGAGCAACGTGGCACTTCTGGAGACCGACGAC 150
|||||
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTyrPyrCysArg 67
|||||
151 GACTCTGCTATAGAGCACTCAGAGCAGAGATGGCCAACTGGTCCGCCCA 200
|||||
67 scysPheProCysCysArgGlySerGlyLysSerAsnValGlyAlaSerG 84
|||||
201 CTGCTTCCCTGCTGCTGAGGGAGGTGGCAAGCAACGTGGCGCTCTG 250
|||||
84 LysAspHisAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100
|||||
251 GAGACCCAGACAGACTGCTGATGAACACCTCAGCAACAGATGGGCAAG 300
|||||
101 TTPCysGlyHisCysPheProCysCysArgGlySerGlyLysSerLysVa 117
|||||
301 TGGTGTGCGACGTGCTCCCTCTGACAGGGAGAGCGGAGACAGACAGGT 350
|||||
117 LglValArgLysArgTyrArgAspSerAlaPheMetGluProArgTyrH 134
|||||
351 GGGCGCTGGGGAGACATACATACATACAGTCCCTCTCAGGAGCCAGGACC 400
|||||
134 LsValArgGlyGlnAspLeuAspLysLeuHisValAlaIleTrrPrgLy 150
|||||
401 ACGTCCGTGGAGAAAGATCTGGACAACTCCACAGACGTGCTGGGGGT 450
|||||
151 LysValProArgLysAspLeuIleValIleMetLeuArgAspThrAspValAs 167
|||||
451 AAAGTCCCGCAAGAGATCTCATCTGATGCTCAGGAGCACTGAGTGA 500
|||||
167 nLysLysAspLysGlnLysArgThrAlaLeuHisValAlaSerAlaAsn 184
|||||
501 CAGAAAGAGCAAGCAAAAGAGAGACTGCTACATCTGCGCTGCCAATG 550
|||||
184 LysAsnSerGluValValLysLeuLeuLeuAspArgArgCysGlnLeuAsn 200
|||||
551 GGAATTCAGAAAGTAGTAAACTCTGCTGAGACAGAGATGTCACCTTAAT 600
|||||
201 ValLeuAspAsnLysLysArgThrAlaLeuIleLysValValGlnCysG 217
|||||
601 GTCTTTCAGACAAAGAAAGAGACAGCTGATGAAGGCCGTACAAATCCA 650
|||||
217 nGluAspArgLysAlaLeuMetLeuLeuGlnHisGlyThrAspProAsn 234
|||||
651 GGAAGATGATGTGGGTAAATGTCGAGACACATGCGCTATATCAAAAT 700
|||||
234 LeProAspGlnLysGlyAsnThrThrLeuHisTyrAlaIleTyrAsnGlu 250
|||||
701 TTCACATAGATGAGAAATACACATCTGACACGCTATCATTAATGAAT 750
|||||
251 AspLysLeuMetAlaLysAlaLeuLeuLeuTyrGlyAlaAspIleGluSe 267
|||||
751 GATTAATTTAATGGCCAAAGCACTGCTTATATGATGCTGATATCGAATC 800
|||||
267 rLysAsnLysHisGlyLeuThrProLeuLeuLeuGlyValHisGlnGln 284
|||||
801 AAAAACAAGACATGGCCTCACACCACTGTACTGTGATACATAGCAAA 850
|||||
284 ysgLlnGlnValValLysPheLeuIleLysLysLysAlaAsnLeuAsnAla 300
|||||
851 AACAGCAAGTCGTGAAATTTTATATCAAGAAAGCAAAATTTAAATGCA 900
|||||
301 LeuAspArgTyrGlyArgThrAlaLeuIleLeuAlaValLysCysGlySe 317
|||||
901 CTGGATAGATATGAAAGAGACTGCTCATATCTTGGTATGTTGTGATC 950
|||||
317 rAlaSerIleValSerLeuLeuLeuGlnGlnAsnIleAspValSerSerG 334
|||||
951 AGCAATATATGTCAGCCTTCTACTTGAGCAAAATATGATGATATCTTCTC 1000
|||||
334 1nAspLeuSerGlyGlnThrAlaArgLutryAlaValSerSerHis 350
|||||
1001 AAGATCTATGAGACAGCGCCAGAGATGCTGTTCTTCAAGCATCAT 1050
|||||
351 HisValIleCysGlnLeuLeuSerAspTyrLysGlnLysGlnMetLeu 367
|||||
1051 CATGTAATTTGCCAGTTACTTCTTGACTCAAGAAAGAAAGATGCTAAA 1100
|||||
367 sIleSerSerGluAsnSerAsnProGluAsnValSerArgThrArgAsn 384
|||||
1101 AATCTCTCTGAAACAGCAATCCAGAAATGCTCAAGAACAGAAATA 1150
|||||
```

384 ys 384  
11  
1151 AA 1152

seq\_name: gb\_pat:AX316986

seq\_documentation\_block:

LOCUS AX316986 1590 bp DNA linear PAT 14-DEC-2001

DEFINITION Sequence 323 from Patent WO0190152.

ACCESSION AX316986

VERSION AX316986.1 GI:17900055

KEYWORDS

SOURCE

ORGANISM human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

Patent: WO 0190152-A 323 29-NOV-2001;

CORTEX CORPORATION (US)

FEATURES

Location/Qualifiers

Source

1..1590 /organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT

424 a 403 c 433 g 329 t 1 others

ORIGIN

alignment\_scores:

Quality: 2064.00 Length: 384

Ratio: 5.375 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-810-936-304 x AX316986 ..

Align seq 1/1 to: AX316986 from: 1 to: 1590

1 MetValValGluValAspSerMetProAlaAlaSerSerValIlySPr 17  
|||||  
436 ATGGTGTTGAGGTGATGCATGCCGCTCTCTCTGTGAAGAAGCC 485  
17 ophegIlyleuArgSerIlySerMetGlyIlyStrpCySArgCySPheroc 34  
|||||  
486 ATTGGCTCTAGAGCAAGATGGCAAGTGGCTGCTGCTGCTGCTGCTGCT 535  
34 yScySArgIlySerGlyIlySerAsnValGlyThrSerGlyAspHisAsp 50  
|||||  
536 GCTGCAGGAGAGAGCGCAAGACACACGCGGCACTCTGAGAGACACGAC 585  
51 AspSerAlaMetIlyStrLeuArgSerIlySerMetGlyIlyStrpCySArgH 67  
|||||  
586 GACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 635  
67 sCySPherocCySArgGlySerGlyIlySerAsnValGlyIlyAspSerG 84  
|||||  
636 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 685  
84 IlyAspHisAspSerAlaMetIlyStrLeuArgSerIlySerMetGlyIlyS 100  
|||||  
686 GAGACCAAGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735  
101 TrpCySArgHisCySPherocCySArgGlySerGlyIlySerIlySVal 117  
|||||  
736 TGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 785  
117 IglIyAlaTrpGlyAspIlyTrpAspSerAlaPheMetGluProArgTyrH 134  
|||||  
786 GGCGGCTGGGGAGACTACGATGACAGAGCTTCTATGAGAGCCAGGTACC 835

134 IValArgGlyGluAspLeuAspIlySerLeuHisArgAlaAlaTrpPgly 150  
|||||  
836 AGCTCCGTGGAGAGATCTGGACAGACTCCACAGAGCTGCGTGGGGGGGT 885  
151 IlySValProArgIlyAspLeuIleValMetLeuArgAspTrpAspValAs 167  
|||||  
886 AAAGTCCCAAGAAAGATCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 935  
167 nIlySAspIlySgIlySArgThrAlaLeuHisLeuAlaSerAlaAsnG 184  
936 CAAGAGGACCAAGCAAAAGAGACTGCTCATCTGCTGCTGCTGCTGCTG 985  
184 IlyAsnSerGluValIlyValLeuLeuLeuAspArgArcCySgIlyLeuAsn 200  
|||||  
986 GGAATTCAGAAAGTGTAAACTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1035  
201 ValLeuAspAsnIlySArgThrAlaLeuIleIlySAlaValGInCySgI 217  
|||||  
1036 GTCTTGTGACAAACAAAGAGAGAGCTGCTGATTAAGCCGCTACATGCA 1085  
217 nGluAspGluCySAlaLeuMetLeuGluHisGlyThrAspProAsnI 234  
1086 GGAAGATGAATGCTGCTTAATGTTGCTGGACATGGCACTGATCCAAATA 1135  
234 IeProAspGluTyrGlyAsnThrThrLeuHisTyrAlaIleTyrAsnGlu 250  
|||||  
1136 TTCAGATGAGTATGAGAAATACCACTGCTGCTGCTGCTGCTGCTGCTG 1185  
251 AspIlyLeuMetAlaIlySAlaLeuLeuLeuTyrGlyAlaAspIleGlu 267  
|||||  
1186 GATTAATTAATGAGCCAAAGACAGCTGCTTAATGAGTGAATGATGAAATC 1235  
267 IlySAsnIlySHisGlyLeuThrProLeuLeuGluIlyValHisGluGln 284  
1236 AAAAACAAGCATGGCTCCACACACAGCTTACTGCTGCTGCTGCTGCTG 1285  
284 yGInGlnIlyValIlyPheLeuIleIlySlySAlaAsnLeuAsnAla 300  
|||||  
1286 AACAGCAAGTGTGTAATTTTAAATCAAGAAACCAATTTTAAATCA 1335  
301 LeuAspArgTyrGlyArgThrAlaLeuIleLeuAlaValIlyCySgIly 317  
|||||  
1336 CTGATTAATTAATGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 1385  
317 rAlaSerIleValSerLeuLeuGluGlnAsnIleAspValSerSerG 334  
1386 AACAAGTATAGTCAAGCTTCTACTTGAACAAATATTGATGATCTTCTC 1435  
334 IAspLeuSerGlyGlnThrAlaArgGluTyrAlaValSerSerHis 350  
|||||  
1436 AAGATCTATCTGAGACAGCGCCAGAGAGTATCTGTTCTAGTATCAT 1485  
351 HisValIleCySgIleLeuLeuSerAspTyrIlySgIlySgIlyMetLeu 367  
1486 CATGTAAATTTGCCACTTACTTCTGACTACAAAGAAACAAATCTTAA 1535  
367 sIleSerSerGluAsnSerAsnProGluAsnValSerArgThrArgAsn 384  
1536 AATCTCTTCTGAAACAGCAATCCAGAAATGTCTCAAGAACCAAGATA 1585  
384 ys 384  
11  
1586 AA 1587

seq\_name: gb\_pat:AX316991

seq\_documentation\_block:

LOCUS AX316991 1155 bp DNA linear PAT 14-DEC-2001

DEFINITION Sequence 328 from Patent WO0190152.

ACCESSION AX316991

VERSION AX316991.1 GI:17900058

KEYWORDS

SOURCE human.

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Frudakis,T.N., Reed,S.G., Smith,J.M., Misher,L.E., Dillon,D.C.,
TITLE Reiter,M.W., Wang,A., Skelky,Y.A., Harlocker,S.L. and Day,C.H.
JOURNAL Compositions and methods for the therapy and diagnosis of breast
PATENT: WO 0190152-A 328 29-NOV-2001;
CORIXA CORPORATION (US)
FEATURES
source 1. 1155
Location/Qualifiers
BASE COUNT 346 a 253 c 296 g 260 t
ORIGIN

alignment_scores:
Quality: 2054.00 Length: 384
Ratio: 5.349 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.479

alignment_block:
US-09-810-936-304 x AX316991 ..

Align seg 1/1 to: AX316991 from: 1 to: 1155

1 MetValaIGluValaAspSerMetProAlaAlaSerSerValIysLeuAsp 17
1 ATGGGGGTTGAGTTGATTCATGCCGGCTCTTGTGTAAGAACCC 50
17 OphelIyLeuArgSerIyMetIyIyStrpCySArgCySArgCySArgCyS 34
51 ATTGGTCTCAGAGCAAGATGGGCAAGATGGTCTCCTCTCTCTCTCTCT 100
34 yScyAArgGluSerGlyIySerSerAsnValGlyThrSerGlyAspHisAsp 50
101 GCTCGAGGAGAGCGGCAAGACAGACAGTGGCGCTTGTGAGAACCCAGC 150
51 AspSerAlaMetIyStrIeuarSerIyMetGlyIyStrpCySArgGHI 67
151 GACTCTGCTATGAGACACTCAGAGCAAGATGGCAAGTGGTGGCGCA 200
67 acySpheroCySArgGlyIySerIySerAsnValGlyAlaIaSerG 84
201 CTGCTCTCCCTGCTGCAAGGGAGTGGCAAGCAACGTTGGCTCTCTG 250
84 IyAspHisAspAspSerAlaMetIyStrIeuarSerIyMetGlyIyS 100
251 GAGACACAGACACTCTGTATGAGACACTCAGAACCAAGATGGCAAG 300
101 TTPCySArgHisCySArgCySArgGlyIySerGlyIySerIyVa 117
301 TGGTCTCTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 350
117 IglYAlaIaTrpGlyAspTyraSAspSerAlaPheMetGluProArgTyra 134
351 GGGCGCTGGGAGACTACATGATGAGTGGCTTCTCATGAGCCAGGTACC 400
134 IValaArgGlyIyAspLeuAspIySerIyMetIyStrIeuarSerIyVa 150
401 ACGTCTCGTGGAGAGATCTGGACAGCTCCACAGAGCTCCCTGGTGGGT 450
151 LysValProArgIyAspLeuIleValMetLeuArgAspPthAspValas 167
451 AAAGTCCCCAGAAAGATCTCATCTCATGCTCAGGAGACACTGAGCTGAA 500
167 nLySAspIySArgIyStrIeuarSerIyMetIyStrIeuarSerIyVa 184
501 CAGAGGAGCAAGCAAGAGAGACTGCTCTACATCTGGCTCTGCTCAATG 550
184 IyAsnSerGluValIyLysLeuLeuAspArgArgCySglnLeuAsn 200

```

---

```

|||||
551 GGAATTGAGAAAGTAGTAAACCTCTGTCGACAGACGATGTCACCTTAAT 600
201 ValLeuAspAsnIySArgIyStrIeuarSerIyMetIyStrIeuarSerIyVa 217
601 GTCTCTTACACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 650
217 nGluAspGluCySAlaLeuMetLeuLeuGluIySgIyThrAspProAsn 234
651 GGAAGATGAATGTGCGTAAATGTTGCTGAGACATGAGCATGATCCAAATA 700
234 IeProAspGluTyraGlyAsnThrIeuarSerIyMetIyStrIeuarSerIyVa 250
701 TTCCAGATGAGTATGGAATATACCACTCTGCTACAGCTATCTATATGAA 750
251 AspIySerMetAlaIySAlaLeuLeuLeuIyGlyIyAlaAspIleGluSe 267
751 GATTAATTAATGAGCAAGACAGCTCTTATATGCTGCTGATATGCAATC 800
267 IyAsnIySArgIyLeuThrProLeuLeuGlyIyAlaIySglnLeu 284
801 AAAAACAAGCATGGCTCTCAGACACTGTTACTTGTGATGATGACGCAA 850
284 ySglnGlnValIySPhelIyIySglnIySAlaAspLeuAsnAla 300
851 AACAGCAAGTGTGAATTTTATTAAGAAAAAGCAATTTTAAATGCA 900
301 LeuAspArgTyraGlyIyStrIeuarSerIyMetIyStrIeuarSerIyVa 317
901 CTGATATGATATGAAAGACTGCTCATATCTGCTGATATGTTGTGGATC 950
317 IAlaSerIeuarSerIeuarSerIeuarSerIeuarSerIeuarSerIeuar 334
951 AGCAAGTATGTCAGCTCTTACTTATGCAAAATTTATATATATCTCTCTC 1000
334 IAspLeuSerGlyIyStrIeuarSerIyMetIyStrIeuarSerIyVa 350
1001 AAGATCTATCTGACACAGCGCCAGAGATGATGCTTCTGATGATCAT 1050
351 HisValIleCySglnLeuLeuSerAspTyraIySglnIySglnIyMetLeu 367
1051 CATGTAATTTGCGAGTACTTCTGACTACAAAGAAAAAGATGCTTAA 1100
367 sIleSerSerGluAsnSerProGluAsnValSerIyStrIeuarSerIyVa 384
1101 AATCTCTTGAAGACGACATCCAGAAATGCTTCAAGAACCAAGAAATA 1150
384 ys 384
1151 AA 1152

seq_name: gb_pat:AX106593
seq_documentation_block:
LOCUS AX106593 2000 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 374 from Patent WO0125272.
ACCESSION AX106593
VERSION AX106593.1 GI:13922264
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2000)
AUTHORS Xu,D., Skelky,Y.A., Reed,S.G. and Cheever,M.A.
TITLE Compositions and methods for therapy and diagnosis of prostate
cancer
JOURNAL Patent: WO 0125272-A 374 12-APR-2001;
CORIXA CORPORATION (US)
FEATURES
source 1. 2000
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

```



BASE COUNT 698 a 388 c 489 g 425 t  
ORIGIN

Alignment\_scores: Quality: 2024.00 Length: 376  
Ratio: 5.383 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment block:  
US-09-810-936-304 x AX106593 ..

Align seg 1/1 to: AX106593 from: 1 to: 2000

```

1 MetValaIgluValAspSerMetProAlaIaSerSerValIylsPr 17
|||||
1 ATGGTGGTGGAGGTTGATTCATGCCGGCTGCTCTCTGTGAAGAAGCC 50
17 OphEgLYleuArgSerIysMetGlyLysTrpCysArgCysPheProc 34
|||||
51 ATTGGCTCAGAGCAAGATGGGCAAGTGGTGGCTGGCTTCCCT 100
34 YsCysArgIuSerGlyLysSerAsnValGlyThrSerGlyAspHisAsp 50
|||||
101 GCTCAGAGAGAGCGCAAGACAGAGTGGGACTTCTGGAGACCAGAC 150
51 AspSerAlaMetLysThrLeuArgSerIysMetGlyLysTrpCysArgH1 67
|||||
151 GACTCTCTATGAAGACTCAGAGCAAGATGGGCAAGTGGCTCCCA 200
67 sCysPheProcCysArgGlySerGlyLysSerAsnValGlyIaIaSerG 84
201 CTGCTTCCCTGCTGCGAGGGGAGTGGCAAGACAGTGGGCTTCTG 250
84 LysAspHisAspAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100
|||||
251 GAGACCCAGAGACGCTGCTATGAACACCTCAGAACAAATAGGGAAG 300
101 TrpCysArgHisCysPheProcCysArgGlySerGlyLysSerLysVa 117
301 TGGTGTGCGCCACTGCTCCCTGCTGCAAGGGGAGCGGCAAGACAGT 350
117 IGLValATrpgLYAspTrAspAspSerAlaPheMetGluProArgTrH 134
351 GGGGCGCTTGGGAGACTACGATACAGTCCCTTCATGAGGCCAGGTACC 400
134 IValaIArgGlyLysAspLysAspLysLeuHisArgAlaIaIaTrpGly 150
401 ACGTCCGAGAGAGATCTGGCAAGCTCCACAGAGCTGCTGGTGGGGT 450
151 LysValProArgLysAspLeuIleValMetLeuArgAspTrHAspValAs 167
451 AAAGTCCCAAGAAAGATCTCATCGTCATGCTCAGGAGCACTGACGTAA 500
167 nLysLysAspLysGlnLysArgTrHAlaLeuHisIleuAlaSerIaAsn 184
501 CAAGAGAGCAAGCAAGAAAGAGACTGCTCTACATCTGGCCCTCCCAATG 550
184 LysAsnSerGluValValLysLeuLeuLeuAspArgArgCysGlnLeuAsn 200
551 GGAATTCAGAGTAGTAAACCTCTGCTGACAGAGATGTCAACTTATAT 600
201 ValLeuAspAsnLysLysArgTrHAlaLeuIleLysAlaValaIaGlnCysG 217
601 GTCCCTTGCAACAAAGAGAGAGAGCTCTGATAAAGCGCGTAAATGCCA 650
217 nGluAspGluCysAlaLeuMetLeuLeuGluHisGlyThrAspProAsnI 234
651 GGAAGATGAATGTGCTTAAATGTGCTGGAACATGCGCATCTCCAAATA 700
234 IeProAspGluTrpGlyAsnTrHThrLeuHisTrpAlaIleTrpAsnGlu 250
701 TTCAGAGATGATGAAATACCACTCTGCTACTACGTATCTATATATGAA 750

```

```

251 AspLysLeuMetAlaLysAlaLeuLeuTrpGlyAlaAspIleGluSe 267
|||||
751 GATAAATTAATGGCCAAAGCACTGCTTATATAGGTGATGATTCGAATTC 800
267 rLysAsnLysHisGlyLeuThrProLeuLeuGluValaHisGluGlnL 284
|||||
801 AAAAACAAGCATGGCCTCACACCACACTGTACTGGTGTACATGAGCAAA 850
284 YsGlnGlnValValLysPheLeuIleLysLysLysAlaAsnLeuAsnAla 300
851 AACAGCAAGTCTGTAAATTTTATATCAGSAAAGCAATTTAATATCA 900
301 LeuAspArgTrpGlyArgTrHAlaLeuIleLeuAlaValaCysCysGlySe 317
|||||
901 CTGGATGATATGGAAGAGAGAGCTCTCATATACGTGCTGTGTGTGGATTC 950
317 rAlaSerIleValSerLeuLeuGluGlnAsnIleAspValaSerSerG 334
951 AGCAAGTATAGTCAGCCCTTCTACTTGTAGCAAAATATGATGATCTCTC 1000
334 LAspLeuSerGlyClnThrAlaArgLysTrpAlaValaSerSerHis 350
1001 AAGATCTATCTGGACAGAGCGCCAGAGAGTATGCTTTCTATGATATCAT 1050
351 HisValIleCysGlnLeuLeuSerAspTrpLysGluLysGlnMetLeuLys 367
1051 CATGTAATTTGGCAGTTTACTTTCGACTACAAAGAAACAGATGCTTAA 1100
367 sIleSerSerGluAsnSerAsnProGlu 376
1101 AATCTCTCTGAAACACCATTCAGAA 1128
seq_name: gb_pat:AX140884
seq_documentation_block:
LOCUS AX140884 2000 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 374 from Patent WO0134802.
ACCESSION AX140884
VERSION AX140884.1 GI:14280987
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2000)
Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
Reed,S.G., Kales,M.D., Retter,M.W., Stoik,J.A., Day,C.H.,
Skeiky,Y.A. and Wang,A.
Compositions and methods for the therapy and diagnosis of prostate
cancer
Patent: WO 0134802-A 374 17-MAY-2001;
CORIXA CORPORATION (US)
location/Qualifiers
source 1..2000
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 698 a 388 c 489 g 425 t
ORIGIN
Alignment_scores: Quality: 2024.00 Length: 376
Ratio: 5.383 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
Alignment block:
US-09-810-936-304 x AX140884 ..
Align seg 1/1 to: AX140884 from: 1 to: 2000
1 MetValaIgluValAspSerMetProAlaIaSerSerValIylsPr 17
|||||

```

```

1 ATGCGTGTGAGGTGATTCATGCGCGGCTGCTCTTCTGTGAAAGACC 50
17 ophegilyleuargserlysmetglylystrpcyscysargcyspheproc 34
   |||||||
51 ATTTGGTCTCAGAGCAAGATGGCAAGTGGTGGCGTTCGCTTCCCT 100
34 yscysargglyuserglylysserasnvalglythrserglyaspheasp 50
   |||||||
101 GCTGCAGGAGAGCGGCAAGACCAACGTGGGACTTCTGGAGACCAGAC 150
51 AspserAlametystrhleuargserlysmetglylystrpcysarghi 67
   |||||||
151 GACTCTGCTATGAGACACTCAGAGCAAGATGGCAAGTGGTCCGCCA 200
67 scyspheprocyscysargglyserglylysserasnvalglylaserg 84
   |||||||
201 CTGCTTCCCTGCTGAGGGGAGTGGCCAAAGCAACGTGGGCGCTTCTG 250
84 lyaspheAspserAlametystrhleuargasnlysmetglylys 100
   |||||||
251 GAGACCAAGACGACTCTGCTATGAGACACTCAGACACAGATGGGCAAG 300
101 Trpcyscysargglyserglylysserasnvalglylaserg 117
   |||||||
301 TGGTGTGCACTGCTTCCCTGCTGCAAGGGGAGCGGCAAGACAGAGT 350
117 lglyAlaTrpGlyAspTrpAspAspserAlaBhmetglubroargtyrh 134
   |||||||
351 GGGCGCTTGGGGAGACTACGATGACAGTCCCTTCATGAGCCCAAGTACC 400
134 lsvAlaTrpGlylyuaspheAspserlysmetglylystrpcysarggly 150
   |||||||
401 ACGTCCCTGGAGAAAGATCTGGACAGCTCCAGAGAGCTGCTGGGGGT 450
151 LysValProarglyAspLeuileValmetleuargspthrAspValas 167
   |||||||
451 AAACTCCCGAGAAAGATCTCATGTCATGTCAGGGAGACCTGACGTGAA 500
167 nlyAspAspLySGlnlysnrthrAlaLeuileuAlaSerAlaasn 184
   |||||||
501 CAAGAGAGACAGCAAAAGAGAGCTGCTCATGTCAGGGAGACCTGCAATG 550
184 lyasnserglyuAlaValysleuLeuLeuAspArgArgCysGlnleuasn 200
   |||||||
551 GGAATTCAGAAAGTAAACTCTCTGACAGACAGATGTCACATTAT 600
201 ValLeuaspasnlylysargthrAlaLeuilelysalavalGlnCysG 217
   |||||||
601 GTCCCTGACAAACAAAAGAGAGACCTCGATTAAGGCCGTAACAATGCCA 650
217 ngluAspGlyCysAlaLeuLeuLeuGlnHISglyThrAspProasn 234
   |||||||
651 GGAAGATGAAGTGCCTGTTATGTTGCTGGAGACATGGCACTGATCCAATA 700
234 leProAspGlyTrpGlyAsnThrThrLeuHISThrAlaIleArgasn 250
   |||||||
701 TTCCAGATGAGTATGGAATATCCACTCTGACGCTATCTATATATATAA 750
251 AspLyseuMetAlaLysAlaLeuLeuLeuTrpGlyAlaAspIleGlyse 267
   |||||||
751 GATTAATTAATGGCCAAAGCACTGCTTATATATGAGTGAATATGATATC 800
267 rlyAsnlysnHISglyLeuThrProLeuLeuGlyValHISGln 284
   |||||||
801 AAAAACAAGATGGCTCAGACACTGTTACTGTTGCTACATGGAGCAAA 850
284 ysgInglValValysPheLeuilelysnlysnlysnleuasn 300
   |||||||
851 AACAGCAAGTCTGAAATTTTATATCAGAAAAAGCAATTAATATCA 900
301 leuAspArgTrpGlyArgThrAlaLeuileuAlaValCysCysGlySe 317
   |||||||
901 CTGATATGATATGGAAGAGTCTCTCATATCTTGTCTGTATGTGGATC 950

```

```

317 rAlaSerIleValSerleuLeuGlnGlnAsnIleAspValSerSerg 334
   |||||||
951 AGCAAGTATATGTCAGGCTTCTACTTGAAGCAAAATATATGATGATCTTCC 1000
334 lAspLeuSerGlyGlnThrAlaArgIleuValAlaSerSerHISHis 350
   |||||||
1001 AAGATCTATCTGACAGAGGCGCAGAGAGTATGCTGTTCTTACTCATATCA 1050
351 HIsValIleCysGlnleuLeuSerAspTrpLySGlnlysnleuLeu 367
   |||||||
1051 CATGTAATTTGGCAGTACTTCTTCTGACTACAAAGAAAAACAGATGCTAAA 1100
367 sIleSerGlyuAsnSerAsnProGlu 376
   |||||||
1101 AATCTCTCTGAAACAGCATCCAGAA 1128

seq_name: gp_pat:AX200744

seq_documentation_block:
LOCUS      AX200744                2000 bp      DNA      linear      PAT 29-AUG-2001
DEFINITION Sequence 374 from Patent WO0151633.
ACCESSION  AX200744
VERSION    AX200744.1  GI:15390637
KEYWORDS   .
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 2000)
            Xu,J., Dillong,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
            Compositions and methods for the therapy and diagnosis of prostate
            cancer. WO 0151633-A 374 19-JUL-2001;
            CORLIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source           1..2000
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        698 a      388 c      489 g      425 t
ORIGIN
1 MethylValGluValAspSerMetProAlaAspSerValysLysr 17
   |||||||
1 ATGCGTGTGAGGTGATTCATGCGCGGCTGCTCTTCTGTGAAAGACC 50
17 ophegilyleuargserlysmetglylystrpcyscysargcyspheproc 34
   |||||||
51 ATTTGGTCTCAGAGCAAGATGGCAAGTGGTGGCGTTCGCTTCCCT 100
34 yscysargglyuserglylysserasnvalglythrserglyaspheasp 50
   |||||||
101 GCTGCAGGAGAGCGGCAAGACCAACGTGGGACTTCTGGAGACCAGAC 150
51 AspserAlametystrhleuargserlysmetglylystrpcysarghi 67
   |||||||
151 GACTCTGCTATGAGACACTCAGAGCAAGATGGGCAAGTGGTCCGCCA 200
67 scyspheprocyscysargglyserglylysserasnvalglylaserg 84
   |||||||

alignment_scores:
Quality: 2024.00      Length: 376
Ratio: 5.383          Gaps: 0
Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
us-09-810-936-304 x AX200744 ..
Align seg 1/1 to: AX200744 from: 1 to: 2000
1 MethylValGluValAspSerMetProAlaAspSerValysLysr 17
   |||||||
1 ATGCGTGTGAGGTGATTCATGCGCGGCTGCTCTTCTGTGAAAGACC 50
17 ophegilyleuargserlysmetglylystrpcyscysargcyspheproc 34
   |||||||
51 ATTTGGTCTCAGAGCAAGATGGCAAGTGGTGGCGTTCGCTTCCCT 100
34 yscysargglyuserglylysserasnvalglythrserglyaspheasp 50
   |||||||
101 GCTGCAGGAGAGCGGCAAGACCAACGTGGGACTTCTGGAGACCAGAC 150
51 AspserAlametystrhleuargserlysmetglylystrpcysarghi 67
   |||||||
151 GACTCTGCTATGAGACACTCAGAGCAAGATGGGCAAGTGGTCCGCCA 200
67 scyspheprocyscysargglyserglylysserasnvalglylaserg 84
   |||||||

```

```

201 CTGCTTCCCTGCTGCAGGGGAGTGGCAAGACACGTGGCGCTTCG 250
84  LyspHisAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100
251 GAGACCCAGCGACTCTGCTATGAGACACTCAGAACAGATGGGAG 300
101 TTPCysHisCysPheProCysCysArgLysSerGlyLysSerLysVa 117
301 TGGTGCTGGCCACTGCTTCCCTGCTGCAGGGGAGCGCAAGAGCAAGGT 350
117 LGLValTrpGlyAspTyrAspSerAlaPheMetGluProArgTyrH 134
351 GGGGGCTGGGGAGACTAGCATGACAGTGCCTCATGGAGCCAGTACC 400
134 LsValArgGlyLysAspLysPheLysHisArgAlaAlaTrpTrpGly 150
401 ACGTCGTGGAGAGATCTGGACAAGCTCCACAGAGCTGCTGGTGGGT 450
151 LysValProArgLysAspLeuIleValMetLeuArgAspThrAspValas 167
451 AAAGTCCCCGAAAGGATCTCATGCTCATGCTCAGGGCACTGACGTGAA 500
167 nLysLysAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsnG 184
501 CAGAGAGCAACAGCAAAAGAGACTGCTTACATCTGGCCTTCGCCAATG 550
184 LysAsnSerGluValLysLeuLeuAspArgArgCysGlnLeuAsn 200
551 GGAATTCAGAGTAGTAAACCTCGTGGACAGACGATGTCACACTTAAT 600
201 ValLeuAspAsnLysLysArgThrAlaLeuIleLysAlaValGlnCysG 217
601 GTCCCTGACAAACAAAGAGAGACTGCTGATTAAGCGCCACAAATGCCA 650
217 nGluAspGluCysAlaLeuMetLeuLysGlnHisGlyThrAspProAsn 234
651 GGAAGATGAATGGCTTAACTTCTGCAACATGGCACTGATCCAAATA 700
234 LeProAspGluTyrGluLysAsnThrThrLeuHisTyrAlaIleTyrAsnGlu 250
701 TTCACAGTAGATGGAAATACCACTGCACCTACGCTATCTAATAATGAA 750
251 AspLysLeuMetAlaLysAlaLeuLeuTyrGlyAlaAspIleGluSe 267
751 GATTAATTAATGGCCAAAGCACTGCTTATATGGTGTGATATCGAATC 800
267 LysAsnLysHisGlyLeuThrProLeuLeuLysGlyValHisGlnGln 284
801 AAAAACAACAGATGGCTTACACACACTGTTACTGGTGTACATGAGCAAA 850
284 YsGlnGlnValLysPheLeuIleLysLysLysAlaAsnLeuAsnAla 300
851 AACGCAAGTCGTGAATTTTAACTCAAGAAAAACGCAATTAAATGCA 900
301 LeuAspArgTyrGlyArgThrAlaLeuIleLeuAlaValCysCysGlySe 317
901 CTGGATATGATATGAAAGACCTGCTCATCTGCTGATGTGTGTGATC 950
317 rAlaSerIleValSerLeuLeuLysGlnGlnAsnIleAspValSerSerG 334
951 AGCAAGATATGTCAGCTTCTACTGACCAAAATATTGATGTATCTTCTC 1000
334 LnsAspLeuSerGlyGlnThrAlaArgLysTyrAlaValSerSerHisHis 350
1001 AAGGTCTATCTGGACAGCGCCAGAGAGTATGCTGTTTCTAGACATCAT 1050
351 HisValIleCysGlnLeuLeuSerAspTyrLysGlnLysGlnMetLeuLys 367
1051 CAGGTAAATTTGCCAGTTACTTCTGACTACAAAGAAAACAGATGCTTAA 1100
367 sLleSerSerGluAsnSerAsnProGlu 376
1101 AATCTCTTCTGAAAAACAGCAATCCAGAA 1128

```

```

seq_name: gb_pat:AX267400
seq_documentation_block:
LOCUS AX267400 2000 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 374 from Patent WO01/3032.
ACCESSION AX267400
VERSION AX267400.1 GI:16516170
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (Stiles)
Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
Kalos,M.D., Fanger,G.R., Reller,M.W., Stolk,V.A., Day,C.H.,
Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skelky,Y.A., Hepler,W.T.
and Henderson,R.A.
Compositions and methods for the therapy and diagnosis of prostate
cancer
Patent: WO 01/3032-A 374 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
location/Qualifiers
source 1..2000
/db_xref="taxon:9606"
BASE COUNT 698 a 388 c 489 g 425 t
ORIGIN
alignment_scores:
Quality: 2024.00 Length: 376
Ratio: 5.383 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-810-936-304 x AX267400 ..
Align seg 1/1 to: AX267400 from: 1 to: 2000
1 MetValAlaGluValAspSerMetProAlaAlaSerValLysLysPr 17
1 ATGGGGTGTGAGGTTGATTCATCCGCGCTCTTCGTGTGAAGAGCC 50
17 orpHeGlyLeuArgSerLysMetGlyLysTrpCysArgCysPheProC 34
51 ATTGTGTCAGAGCAAGATGGCAAGTGTGCTCCCTCTCCCTCCCT 100
34 YsCysArgGluSerGlyLysSerAsnValGlyThrSerGlyAspHisasp 50
101 GCTGCAGGAGAGCGCGCAAGACACTGGCGCACTTCTGAGACACAGAC 150
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTrpCysArgH 67
151 GACTGCGTATGAAAGACACCTCAGGAGCAAGATGGCAATGGTGGCGCA 200
67 scysPheProCysCysArgGlySerGlyLysSerAsnValGlyAlaSerG 84
201 CTGCTTCCCTGCTGCAGGGGAGTGGCAAGAGCAAGTGGCGCTTCG 250
84 LyspHisAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100
251 GAGACCCAGCGACTCTGCTATGAGACACTCAGAACAGATGGGAG 300
101 TTPCysHisCysPheProCysCysArgLysSerGlyLysSerLysVa 117
301 TGGTGCTGGCCACTGCTTCCCTGCTGCAGGGGAGCGCAAGAGCAAGGT 350
117 LGLValTrpGlyAspTyrAspSerAlaPheMetGluProArgTyrH 134
351 GGGGGCTGGGGAGACTAGCATGACAGTGCCTCATGGAGCCAGTACC 400
134 LsValArgGlyLysAspLysPheLysHisArgAlaAlaTrpTrpGly 150

```

```
|||||
401 ACGTCCGTGGAGNAGTCTGGACAGCTCCACAGAGCTGCTGGTGGGCT 450
151 LysValProArgLysAspLeuIleValMetLeuArgAspThrAspValAs 167
451 AAAGTCCCGCAGAAAGATCTCATCTCGATGCTCAGGACACTGACGTGAA 500
167 nLysLysAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsn 184
501 CAAAGAGACACAGCAAAAGAGGACTGCTCTACATCTGGCCCTCCCATG 550
184 LysAsnSerGluValValLysLeuLeuLeuAspArgArgCysGlnLeuAsn 200
551 GGAATTCAGAAAGTAGTAAGAACTCTCTCTGGAGACAGATGCAACTTAT 600
201 ValLeuAspAsnLysLysArgThrAlaLeuIleLysAlaValGlnCysG 217
601 GTCCCTGACAAACAAAAGAGACAGCTGTGATTAAGCCGTCATCATGCCA 650
217 nGluAspGluCysAlaLeuMetLeuLeuGlnHisGlyThrAspProAsn 234
651 GGAAGATGATGTCCGTATATGCTGTCTGAGACATGCGCATGTCACATA 700
234 LeProAspGluTrpGlyAsnThrThrLeuHisTrpAlaIleTrpAsnGlu 250
701 TTCACAGATGAGTATGGAAATACCACTCTGACTACGCTATATTAATGAA 750
251 AspLysLeuMetAlaLysAlaLeuLeuLeuTrpGlyAlaAspIleGluSe 267
751 GATTAATTAATGGCCAAAGCAGCTGCTTATATGGTGGTGATATGCAATC 800
267 rLysAsnLysHisGlyLeuThrProLeuLeuGlnLysValHisGlnGln 284
801 AAAAACAAGATAGCCCTCACACCAGCTGTACTGTGTGACATGGACAA 850
284 YsgGlnValValLysPheLeuIleLysLysAlaAsnLeuAsnAla 300
851 AACAGCAAGTCGTGAATTTTATCATCAAGAAAAGCGAATTTAAATGCA 900
301 LeuAspArgTrpGlyArgThrAlaLeuIleLeuAlaValCysCysGlySe 317
901 CTGGATGATATGGAAGAGACTGCTCATACATCTGCTATATGTGTGGATC 950
317 rAlaSerIleValSerLeuLeuGlnGlnAsnIleAspValSerSerG 334
951 AGCAAGATATAGTCAGCCCTCTACTTACGCAAAATATGTGATCTTCTC 1000
334 LnaAspLeuSerGlnThrAlaArgGluTrpAlaValSerSerHisHis 350
1001 AAGATCTATCTGGACAGACGGCCAGAGATGCTGTTTCTATGTCATCAT 1050
351 HisValIleCysGlnLeuLeuSerAspTrpLysGlnLysGlnMetLeuLy 367
1051 CATGTAATTTGGCAGTACTTCTGTACTGCAAGAAAGAAACAGATGCTAAA 1100
367 sIleSerSerGlnAsnSerAsnProGlu 376
1101 AATCTCTCTGAAAAACAGCAATCCAGAA 1128

seq_name: gb_pat:AX282957
seq_documentation_block:
LOCUS AX282957 2000 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 6 from Patent WO0175171.
ACCESSION AX282957
VERSION AX282957.1 GI:16609897
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (sites) Houghton,R.L., Dillon,D.C., Molesh,D.A., Xu,J., Zehentner,B. and
AUTHORS
```

```
Persing,D.H.
Methods, compositions and kits for the detection and monitoring of
breast cancer
Patent: WO 01/75171-A 6 11-OCT-2001:
JOURNAL CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..2000
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 698 a 388 c 489 g 425 t
ORIGIN

alignment_scores:
Quality: 2024.00 Length: 376
Ratio: 5.383 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
us-09-810-936-304 x AX282957 ..
Align seg 1/1 to: AX282957 from: 1 to: 2000

1 MetValValGlnValAspSerMetProAlaAlaSerValLysLysPr 17
|||||
1 ATGGGTGTTGAGGTGATTCATTCATGCGGCTGCTCTCTGTGAAAGAGCC 50
17 opheGlyLeuArgSerLysMetGlyLysTrpCysArgCysPheProC 34
|||||
51 ATTTGGTCTCAGAGCAGATGGCAAGTGGTGGTGGCTTCCCTTCCCT 100
34 yscYsArgGlnSerGlyLysSerAsnValGlyThrSerGlyAspHisAsp 50
|||||
101 GCTTCAGAGAGAGCGCAGACAGCAACGTGGCCTTCTGAGACACAGC 150
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTrpCysArgH 67
|||||
151 GACTCTGCTATGAAACACTCAGAGCAAGATGGCAAGTGTGTCGCCCA 200
67 scYsPheProCysCysArgGlySerGlyLysSerAsnValGlyAlaSerG 84
201 CTGCTTCCCTGCTGCTGAGGGGAGTGGCAAGCAGCTGGGCGCTTCTG 250
84 LysAspHisAspAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100
251 GAGACCAAGCAGACTGCTGTAAGACACTCAGAGACATGAGACATGGGCA 300
101 TrpCysCysHisCysPheProCysCysArgGlySerGlyLysSerLysVa 117
|||||
301 TGGTGTGCGCACAGTCTTCCCTGCTGACAGGGGAGCGGCAAGACAGGT 350
117 lGlyAlaTrpGlyAspTrpAspAspSerAlaPheMetGlnProArgTrpH 134
|||||
351 GGGGGCTTGGGGAGACTACGATGACAGTGCTTCTTCTGAGAGCCCAAGT 400
134 lValAlaArgGlyGlnAspLeuAspLysLeuHisArgAlaAlaTrpTrpGly 150
401 ACGTCCGTGGAGAAATCTGGACAAAGCTCCACAGAGCTGCTGTGGTGG 450
151 LysValProArgLysAspLeuIleValMetLeuArgAspThrAspValAs 167
451 AAAGTCCCGCAGAAAGATCTCATCTCGATGCTCAGGACACTGACGTGAA 500
167 nLysLysAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsn 184
501 CAAAGAGACACAGCAAAAGAGGACTGCTCTACATCTGGCCCTCCCATG 550
184 LysAsnSerGluValValLysLeuLeuLeuAspArgArgCysGlnLeuAsn 200
551 GGAATTCAGAAAGTAGTAAGAACTCTCTCTGGAGACAGATGCAACTTAT 600
201 ValLeuAspAsnLysLysArgThrAlaLeuIleLysAlaValGlnCysG 217
|||||
```

```
601 GTCTTGACAACAAAAAGAGACAGCTCGATAAAGCCGTACATGCCA 650
217 ngluapbplucysalaleuMetleuLeuGluHISGLYthcAspProasnI 234
|||||
651 GGAAGATGAAATGTGGCTTAATGTGCTGTGAAACATGGCACATGCATCAAAAT 700
234 leProaspGluTYrGLYasnThrThrLeuHISYrAlaIleTYrAsnGlu 250
|||||
701 TTCAGATGAGTATGGAAATACCACTGCACTGCACTGCACTGCTATTAATGAA 750
251 AspLysLeuMetAlaLysAlaLeuLeuTYrGLYAlaAspIleGluSe 267
|||||
751 GATTAATTAATGGCCAAAGCACTGCTTATATGCTGTATATCGAATC 800
267 rLysAsnLysHISGLYleuThrProLeuLeuGluYValHISGLUGln 284
|||||
801 AAAAACAAGCATGGCCCTCACACACTGTTACTTGCTGTACATGAGCAAA 850
284 YSGInGlnValLysPheLeuIleLysLysAlaAsnLeuAsnAla 300
|||||
851 AACGCAAGTCGTGAATTTTATATCAAGAAAAAGCAATTTAATGCA 900
301 leuAspArgTYrGLYArgThrAlaLeuIleLeuAlaValLysCysGLYSe 317
|||||
901 CTGGATGAGTATGGAAAGACTGCTCCACTGCTGCTGATGTGTGGATC 950
317 rAlaSerIleValSerLeuLeuGluGlnAsnIleAspValSerSerg 334
|||||
951 AGCAAGTATAGTCAGCTTCTACTTGAGCAAAATATGATGTATCTTCTC 1000
334 lnaAspLeuSergLYnThrAlaArgGLYrTYrAlaValSerSergHISHS 350
|||||
1001 AAGATCATATCTGGACAGCGCCAGAGATAGCTGTTCAGTCATCAT 1050
351 HIsValIleCysGlnLeuLeuSergAspTYrLysGLYLysGlnMetLeu 367
1051 CATGTAATTTGCCAGTTACTTCTGTACACAGAAACAGATGACTAAT 1100
367 sTlSerSergLysAsnSeraProGlu 376
|||||
1101 AATCTCTTCTGAAACAGCAATCCAGAA 1128
seq_name: gb_pat:AX316965
seq_documentation_block:
LOCUS AX316965 2000 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 302 from Patent WO0190152.
ACCESSION AX316965
VERSION AX316965.1 GI:17900044
KEYWORDS
SOURCE
ORGANISM human.
REFERENCE Homo sapiens
AUTHORS Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)
TITLE Rutter M.W., Wang A., Skelky Y.A., Harlocker, S.L. and Day, C.H.
COMPOSITIONS and methods for the therapy and diagnosis of breast
cancer
JOURNAL Patent: WO 0190152-A 302 29-NOV-2001;
CORIXA CORPORATION (US)
FEATURES
SOURCE 1. 2000
Location/Qualifiers
BASE COUNT 698 a 388 c 489 g 425 t
ORIGIN
alignment_scores:
Quality: 2024.00 Length: 376
Ratio: 5.383 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
us-09-810-936-304 x AX316965
Align seg 1/1 to: AX316965 from: 1 to: 2000
1 MetValValGluValAspSeraProAlaIleSerSerValLysLysPr 17
|||||
1 ATGATGATGAGTGTGATTCATCCAGCTGCTCTTCTGTGTAAAGACC 50
17 opheGlyLeuArgSerLysMetGlyLysTYrCysCysArgCysPheProC 34
|||||
51 ATTTGCTCTCAGGAGCAAGATGGCCAAAGTGCTGCTGCTCTCTCTCCCT 100
34 YscYsArgLysSergLysSeraValGlyThrSergLysAspHisAsp 50
|||||
101 GCTGCAGGGAGAGCGGCAAGAGCAACGTTGAGCACTTGGAATCCAGCAC 150
51 AspSeraIleMetLysThrLeuArgSerLysMetGlyLysTYrCysArgH 67
|||||
151 GACTCTGCTATGAAACACTCAGAGCAACATGGCAATGGTGGCGCCA 200
67 sCysPheProCysCysArgGlySergLYnLysSeraValGlyAlaSerg 84
|||||
201 CTGGTTCCTCCTGCTGAGGGAGAGTGGCAAGCAACGTGGGCGCTTCTG 250
84 lYsPheHisAspSeraIleMetLysThrLeuArgAsnLysMetLys 100
|||||
251 GAGACACAGCAGCACTGCTATGAAACACTCAGAAACAGATGGGCAAG 300
101 TrpCysCysHISCysPheProCysCysArgGLYSerGLYSerLysVal 117
|||||
301 TGATGTCGACACCTGCTCCCTGCTGCAAGGGGAGGGCAAGCAAGAGT 350
117 lGlyAlaIleTrpGLYAspTYrAspAspSeraIlePheMetGluProArgTYrH 134
|||||
351 GGGGCTGGGGAGACTACGATGACAGAGTCCCTTATGATGAGCCAGAGTACC 400
134 lSValAlaArgGLYlAspLeuAspLysLeuHISArgAlaIleTrpTYrGLY 150
|||||
401 ACGTCCCTGGAGAGATCTGGCAAGCTCCACAGAGCTCCTGGGGGGT 450
151 LysValProArgLysAspLeuIleValMetLeuArgAspThrAspValAs 167
|||||
451 AAAGTCCCGCAAGAGATCTCATGCTCATGCTCAGGAGACATGACGTGAA 500
167 nLysLysAspLysGlnLysArgTYrAlaLeuHISLeuAlaSerAlaAsn 184
|||||
501 CAAAGAGCAACAAAAAGAGAGACTGCTATCATCTGGCTCTGCCAATG 550
184 lYsAsnSergLysValLysLysLeuLeuAspArgArgCysGlnLeuAsn 200
|||||
551 GGAATTCAGAAAGTAAACATCTCTGCTGACACAGCATGTCACTTAAT 600
201 ValLeuAspAsnLysLysArgTYrAlaLeuIleLysAlaValGlnCysGL 217
|||||
601 GTCTTCACACAAAAAGAGAGCAAGCTCTGATTAAGGCCGTCAATGCA 650
217 ngluAspGLYcysalaleuMetleuLeuGluHISGLYThrAspProasnI 234
|||||
651 GGAAGATGAGTGTGCTTAATGTGCTGGAACATGGACATGCATCAATTA 700
234 leProaspGluTYrGLYasnThrThrLeuHISYrAlaIleTYrAsnGlu 250
|||||
701 TTCAGATGAGTATGGAAATACCACTGCACTGCACTGCTATTAATGAA 750
251 AspLysLeuMetAlaLysAlaLeuLeuTYrGLYAlaAspIleGluSe 267
|||||
751 GATTAATTAATGGCCAAAGCACTGCTTATATGCTGTATATCGAATC 800
267 rLysAsnLysHISGLYleuThrProLeuLeuGluYValHISGLUGln 284
|||||
801 AAAAACAAGCATGGCCCTCACACACTGTTACTTGCTGTACATGAGCAAA 850
```

284 ysgInglInValIalLysPheLeuIleLysLysLysAlaLeuAsnAla 300  
|||||  
851 AACGCAAGTCGTGAATTTTATCAAGAAAAAGCAATTTAAATGCA 900  
301 LeuAspArgTyrGlyArgThrAlaLeuIleLeuAlaValCysCysGlySe 317  
|||||  
901 CTGGATAGATATGAGAGAGCTGCTCATCTTCTGCTGATGTTGGATC 950  
317 rAlaSerIleValSerLeuLeuGluGlnAsnIleAspValSerSerg 334  
|||||  
951 AGCAAGATATGACAGCTTCTACTTGAGCAAAATATGATATCTTCTC 1000  
334 InAspLeuSerGlyGlnThrAlaArgGluTyrAlaValSerSerHisHis 350  
|||||  
1001 AAGATCTATCTGGACGACGCGCAGAGATGCTGTTCTAGTCATCAT 1050  
351 HisValIleCysGlnLeuLeuSerAspTyrLysGluLysGlnMetLeuIly 367  
|||||  
1051 CATGTAATTTGCCAGTACTTCTTGACTACAAAGAAAAACAGATGCTAAA 1100  
367 sIleSerSergLysAsnSerAsnProGlu 376  
|||||  
1101 AATCTCTTCTGAAAAACAGCAATCCAGAA 1128  
seq\_name: gb\_pat:AX106594  
seq\_documentation\_block:  
LOCUS AX106594 2040 bp DNA linear PAT 30-APR-2001  
DEFINITION Sequence 375 from Patent WO0125272.  
ACCESSION AX106594  
VERSION AX106594.1 GI:13922265  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.  
TITLE Compositions and methods for therapy and diagnosis of prostate  
cancer  
JOURNAL Patent: WO 0125272-A 375 12-APR-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source location/Qualifiers  
1..2040  
BASE COUNT 716 a 392 c 500 g 432 t  
ORIGIN  
Alignment\_scores:  
Quality: 2024.00 Length: 376  
Ratio: 5.383 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-810-936-304 x AX106594 ..  
Align seg 1/1 to: AX106594 from: 1 to: 2040  
1 MetValValGluValAspSerMetProAlaIleSerSerValLysLysPr 17  
|||||  
1 ATGGTGGTTGAGGTTCATTCATGCGGCTCCTCTCTGTGAAGAGCC 50  
17 oPheGlyLeuArgSerLysMetGlyLysTrpCysCysArgCysPheProC 34  
|||||  
51 ATTGGTCTCAGAGCAAGATGGCAAGTGGTCTGCCGTGTTCTCCCT 100  
34 yscAsArgGlySerGlyLysSerAsnValGlyThrSergLysAspHisAsp 50  
|||||  
101 GCTCAGAGAGAGCGGCAAGACACAGTGGGCACTTCTGAGACCCAGAC 150

51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTrpCysArgHis 67  
|||||  
151 GACTCTGATATGAGACACTGACGAGCAGATGGCAAGTGGTGGCCGCA 200  
67 scsPheProCysCysArgGlySergLysSerAsnValGlyAlaSerG 84  
|||||  
201 CTGCTTCCCTCTGCTGACAGGGAGAGTGGCAAGACAGCGGGCGTTCTG 250  
84 LysPheHisAspAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100  
|||||  
251 GAGACCCAGCAGACTCTGCTATGAAAGACTGACGAAACAAGATGGGCAAG 300  
101 TrpCysCysHisCysPheProCysCysArgGlySergLysSerLysVal 117  
|||||  
301 TGGTGTGCTGCACTGCTTCCCTGCTGACGAGGAGCGGCAAGACAGAGCT 350  
117 LglValATrpgLysAspTyrAspAspSerAlaPheMetGluProArgTyrH 134  
|||||  
351 GGGCGCTTGGGAGAGCTACGATGACAGTGGCTTCATGAGAGCCAGATACC 400  
134 LsValArgGlyGluAspLeuAspLysLysLeuHisArgAlaIleTrpGly 150  
|||||  
401 ACGTCCGTGAGAGATCTGGACAGGCTCCACAGAGCTGCTGGGGGT 450  
151 LysValProArgLysAspLeuIleValMetLeuArgAspThrAspValAs 167  
|||||  
451 AAGTCCCGCAAGAGATCTCATGCTGCTGACGAGGACACTGACGCTGAA 500  
167 nLysLysAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsnG 184  
|||||  
501 CAAAGAGACAGCAAGAAAGAGAGACTGCTACATCTGGCTCTGCCAATG 550  
184 LysAsnSergLysValLysLeuLeuLeuAspArgArgGlyGlnLeuAsn 200  
|||||  
551 GGAATTCAGAGTACTTAATCTGCTGGAGCAGAGAGATGATACCTTAT 600  
201 ValLeuAspAsnLysLysArgThrAlaLeuIleLysAlaValGlnCysG 217  
|||||  
601 GTCCTTGACAAACAAAAAGAGCAGCTCTATTAAGCCGTCAATGTGCA 650  
217 nGluAspGluCysAlaLeuMetLeuLeuGlnHisGlyThrAspProAsn 234  
|||||  
651 GGAAGATGATGTGGCTTAATGCTTGGAACATGGCACTGATCCAAATTA 700  
234 LsProAspGluTyrGlyLysAsnThrThrLeuHisTyrAlaIleLysAsnG 250  
|||||  
701 TTCCAGATGATATGAAATACACACTGCGACCTACGCTATCATATGAA 750  
251 AspLysLeuMetAlaLysAlaLeuLeuLysGlyAlaAspIleGluSe 267  
|||||  
751 GATTAATTAATGCGCAAGACCTGCTTATATGATGCTGATATCGAATC 800  
267 rLysAsnLysHisGlyLeuThrProLeuLeuGluGlyValHisGluGln 284  
|||||  
801 AAAAACAAGCAGATGGCCCTCACACACAGTACCTGGTGTGACATGAGCAA 850  
284 ysgInglInValIalLysPheLeuIleLysLysLysAlaLeuAsnAla 300  
|||||  
851 AACAGCAAGTCTGAAATTTTATCAAGAAAAACGAATTTAAATGCA 900  
301 LeuAspArgTyrGlyArgThrAlaLeuIleLeuAlaValCysCysGlySe 317  
|||||  
901 CTGGATAGATATGAGAGAGCTGCTCATCTTCTGCTGATGTTGGATC 950  
317 rAlaSerIleValSerLeuLeuGluGlnAsnIleAspValSerSerg 334  
|||||  
951 AGCAAGATATATGACAGCTTCTACTTGAGCAAAATATGATATCTTCTC 1000  
334 InAspLeuSerGlyGlnThrAlaArgGluTyrAlaValSerSerHisHis 350  
|||||  
1001 AAGATCTATCTGACAGAGCGCAGAGATATGCTTCTTCTAGTCATCAT 1050  
351 HisValIleCysGlnLeuLeuSerAspTyrLysGluLysGlnMetLeuIly 367

```
|||||  
1051 CATGTAATTTGGCAAGTTACTTTCTGACTACAGAGAAAACGATGCTAAA 1100  
367 s11eSerSerGIUASnSerASnProGlu 376  
|||||  
1101 AATCTCTCTGTGAAAACAGCAATCCAGAA 1128
```

---







1	ATGGGTGTTGAGGTGATTCATGCCGGCTGCTTCTTGGAAGAAC	50
17	oRhegLYleuAtgSerLYsMetGLYsTTPCySCyArGcYPheProC	34
51	ATTGGCTCAGAGCAAAATGGCAATGGTGGCTGGCTTCCTCCCT	100
34	ySCySArGLuSeRGlyLYsSerAsnValGLYThrSerGLYAspHLSAsp	50
101	GCTGAGGGGAGAGCGCAACAGCAACGTGGCAGCTCTGGGAGACAGAC	150
51	AspSerAlaMetLYsThrLeuArgSerLYsMetGLYsTTPCySArGH	67
151	GACTCTGCTATGTAAGACACTCGAGAGCAAGATGGGCAAGTGGTCCCA	200
67	sCySPheProCySCySArGLYSerGLYLySerAsnValGLYAlaSerG	84
201	CTGCTTCCTCCCTGGTAGGGGAGTGGCAAGCAACGTGGCCTCTTG	250
84	LYsPHisAspAspSerAlaMetLYsThrLeuArgAsnLYsMetGLYs	100
251	GAGACCAACGACGACTGCTATGAAGACACTGAGAACATGGATGGCAAG	300
101	TTPCySCySHisCYsPheProCYsCySArGLYSerGLYLySeriLYsVa	117
301	TGGTGCCTGCACGCTTCCTCCCTGCTGCAGGGGGAGCGCAAGACAGT	350
117	LGlyAlaTPrGLYAspTYrAspAspSerAlaPheMetGLuProArgTYrH	134
351	GGCGCGCTGGGGAGACTACGATGACAGTGGCTCTCATGAGCCCAAGTACC	400
134	LYsValArgGLYsAspLeuAspLYsLeuHisArgAlaAlaTTPTPGLY	150
401	ACGTCGGTAGAGAGATCTGGAACAAGCTCCACAGAGCTGCTGGGGGT	450
151	LYsValProArgLYsAspLeuLevalMetLeuArgAspTYrAspValas	167
451	AAAGTCCCAAGAAAGGATCTCATCTCATGCTGCAGGAGCACTGACGTGA	500
167	nLYsLYsAspLYsGLuLYsArgTYrAlaLeuHisLeuAlaSerAlaasnG	184
501	CAACAAGACACACCAAAAGAGACGCTCTCATCTGGCCTCCCAATG	550
184	LYsAsnSerGLuValLYsLYsLeuLeuLeuAspArgArCYsGLuLeuAsn	200
551	GGAATTCAGAAAGTAGTAAACCTCGTGGAGACAGACGATGTCACATTAA	600
201	ValLeuAspAsnLYsLYsArgTYrThrAlaLeuILEYsAlaValGLuCYsG	217
601	GTCCTTGACACAAAAAGAGACGCTGTGATAAAGGCCGTCACATGCCA	650
217	nGLYAspGLYsAlaLeuMetLeuLeuGLuHisGLYThrAspProAsnI	234
651	GGAAGATGAATGGCGTTAATGTTGGTGAACATGGCAGCTGATCCAAATA	700
234	LeProAspArgLYsTYrGLYAsnThrThrLeuHisTYrAlaILEYsAsnGLu	250
701	TTCAGATGAGTAGTGAATATACCACCTGTGCACATGCGATTTAAAGAA	750
251	AspLYsLeuMetAlaLYsAlaLeuLeuLeuTYrGLYAlaAspILEGlyUse	267
751	GATTAATTAATGAGCCAAAGACGTGGCTTAATGTGCTGATTCGAATC	800
267	TYsAsnLYsHisGLYLeuThrProLeuLeuLeuGLYValHisGLuGLuL	284
801	AAAAAACACCAAGGCGCTCACACCACTGTACTGGTGTACATGAGGAAA	850
284	YsGLuGLuValLYsPheLeuILEYsLYsLYsAlaAspLeuAspAla	300
851	AACGACAGTCGTGAATTTTAATCAAGAAAAAGCAATTTAAATGCA	900
301	LeuAspArgTYrGLYArgTYrThrAlaLeuILEuAlaValCYsCYsGLYse	317

901 CTGATACATATGAGAGACGCTGCTCATACTTGCTGTGATGTGGATC 950

317 TALAser11eValSerLeuLeuLeuGluGlnAsn11eAspValSerSerg 334

|||||

951 AGCAAGTATAGCAGCCTTCTACTTGACCAAAAATTTGATGTATCTTCTC 1000

334 lAspLeuSerg1yGlnThrAlaArg1uTyTAlaValSerSernHis 350

|||||

1001 AAGATCTATCTGGACAGACGGCAGAGAGTATGCTGTTTCTATCATCAT 1050

351 HisVal11eGysGlnLeuLeuSerAspTyrTyrSgluYsglnMetLeu 367

|||||

1051 CAYGTAATTTGCCAGTACTTCTTGACTACAAAGAAAACGATGCTAAA 1100

367 s11eSerSerg1uAsnSernAsnProGluuAsnValSerArgThrArgAsn 384

|||||

1101 AATCTCTCTGTGAAGAACAGCAATCCGAAATAATGTCTCAAGAACAGAAATA 1150

384 ys 384

1151 AA 1152

seq\_name: /SIDS1/gcgcdata/hold-geneseq/geneseq-emb1/NA2000.DAT.AAA06598

seq\_documentation\_block:

ID AAA06598 standard; cDNA; 1155 BP.

AC AAA06598;

DT 13-JUN-2000 (first entry)

DE Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:373.

XX

XX Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:373.

DE

XX Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:373.

XX

XX Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:373.

KM Human; prostate cancer; diagnosis; tumour; gene therapy; detection;

XX immunogenic; cytosolic; vaccine; ss.

XX

OS Homo sapiens.

XX

XX WO200004149-A2.

PN

XX 27-JAN-2000.

PD

XX

XX 14-JUL-1999; 99WO-US15838.

PF

XX

XX 14-JUL-1998; 98US-0115453.

PR 14-JUL-1998; 98US-0116134.

PR 23-SEP-1998; 98US-0159812.

PR 23-SEP-1998; 98US-0159822.

PR 15-JAN-1999; 99US-0232149.

PR 15-JAN-1999; 99US-0232880.

PR 09-APR-1999; 99US-0288946.

XX

XX (CORI-) CORIXA CORP.

PA

XX

PI Dillon DC, Harlocker SL, Yugin J, Xu J, Mitcham JL;

XX

XX WPI; 2000-171268/15.

DR

XX

PT New polypeptide useful for treating and diagnosing prostate cancer

XX comprises an immunogenic portion of prostate tumor protein -

XX

PS Claim 50; Page 222; 263pp; English.

XX

CC The present invention describes isolated polypeptides, comprising an

CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides

CC and polynucleotides encoding them have cytosolic activity and can be

CC used in vaccines and in gene therapy. The polypeptides and

CC polynucleotides encoding them, antigen presenting cells which express

CC the polypeptides, antibodies against the polypeptides and vaccines

CC comprising them can be used for inhibiting the development of prostate

CC cancer in a patient. The polypeptides can be used to generate antibodies

CC or anti-idiotypic antibodies for passive immuno therapy. A portion of

CC the polynucleotides encoding the polypeptides can be used as a probe or

CC to modulate the expression of the polypeptides. AA06241 to AA06691 and  
CC AA082000 to AA082020 represent sequences used in the exemplification of  
the present invention.

XX Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

#### alignment\_scores:

Quality: 2064.00 Length: 384  
Ratio: 5.375 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

#### alignment\_block:

us-09-810-936-304 x AA06598 ..

Align seg 1/1 to: AA06598 from: 1 to: 1155

```

1 MetValValGluValAspSerMetProAlaAlaSerSerValLysLysPr 17
|||||
1 ATGGGTGGTGAAGTATATTCATGCGGCTCTCTCTGTAAGAACGCC 50
17 opheGlyLeuAArgSerLysMetGlyLysTrpCysAArgCysPheProC 34
51 ATTGGTCTCAGAGAGCAAGATGGGCAAGTGGTGGCTGGCTGCCCT 100
34 yscysAArgGlySerGlyLysSerAsnValGlyThrSerGlyAspHis 50
101 GCTGACAGGAGAGGAGGACAGACAGCTTCTGGAGACACAGCAC 150
51 AspSerAlaMetLysThrLeuAArgSerLysMetGlyLysTrpCysAArgH 67
151 GACTCTCTATGAAGACACTCAGAGCAAGATGGGCAAGTGGTGGCT 200
67 scysPheProCysAArgGlySerGlyLysSerAsnValGlyAlaSerG 84
201 CTGCTTCCCTGCTGCGAGGGAGTGGCAAGCAACGTGGGCTTTCTG 250
84 lyaSpHisAspAspSerAlaMetLysThrLeuAArgAsnLysMetGlyLys 100
251 GAGACCCAGACGACTCTGTATGACACTCAGAGCAAGATGGGCAAG 300
101 TrpCysCysHisCysPheProCysAArgGlySerGlyLysSerLysVal 117
301 TGGTGGTGGCACTGCTTCCCTGCTGCAAGGGAGGGGCAAGACAGAGT 350
117 lGlyAlaTrpGlyAspTrpAspAspSerAlaPheMetGluProAArgTr 134
351 GGGGCTTGGGAGACTACGATGACAGTCTTCATGAGGCCAGGTCC 400
134 lSValAArgGlyGluAspLeuAspLysLeuHisAArgAlaAlaTrpTrpGly 150
401 ACGTCCGCGGAGAAAGATCTGGACAGCTCCACAGAGCTCCGCGGGGT 450
151 LysValAlProAArgLysAspLeuLeValMetLeuAArgAspThrAspVal 167
451 AAAGTCCCGCAAGAAAGATCTCATGCTATGCTCAGGAGACAGAGCTAA 500
167 nLysLysAspLysGlnLysAArgThrAlaLeuHisLysLeuAlaSerAla 184
501 CAAGAGAGACAGCAAGAGAGAGACTGCTACATCTGGGCTTGGCAATG 550
184 lYasnSerGluValValLysLeuLeuLeuAspAArgAArgGlySerGln 200
551 GGAATTCAGAGTAGTAAACTCTGCTGGACAGAGAGTGAACCTTAAT 600
201 ValLeuAspAsnLysLysAArgThrAlaLeuLeLysAlaValGlnCysG 217
601 GTCTCTGCAACAACAAAGAGAGAGCTGTGATAAAGGCCGTACATGCA 650
217 nGluAspGlyCysAlaLeuMetLeuGlnHisGlyTrpAspProAsn 234
651 GGAAGATGAATGTCGTTAATGTGCTGGACAGTGGCACTGATCAAAATA 700

```

```

234 lProAspGlyTrpGlyAsnThrThrLeuHisLysTrpAlaAlaLysGln 250
|||||
701 TTCAGATGAGTAGTAAATACACACTGCTACAGCTATGATTAATGA 750
251 AspLysLeuMetAlaLysAlaLeuLeuLeuTrpGlyAlaAspLysG 267
751 GATAAATTTAATGGCCAAAGCAGCTGCTTATATGCTGTGATATGCAATC 800
267 rLysAsnLysHisGlyLeuThrProLeuLeuGlnGlyValHisGlnL 284
801 AAAAACAAGCATGGCTCAGACACACTGTTACTTGTGTTACATGAGCAA 850
284 yscGlnGlnValLysPheLeuLeLysLysAlaAsnLeuAsnAla 300
851 AACAGCAAGTCGTGAATTTTATATCAAGAAAAAGCCAAATTAATGCA 900
301 LeuAspAArgTrpGlyAArgThrAlaLeuLeLysAlaValCysCysGly 317
901 CTGGATATGATATGGAAGAGACTGCTCATATCTTGTGATGATGATCTTC 950
317 rAlaSerLeValSerLeuLeuGlnGlnAsnLeAspValSerSerg 334
951 AGCAAGTATGATCAGCTTCTTACTTACAGCAAAATATGATGATCTTCTC 1000
334 lAspLeuSerGlyGlnThrAlaAArgLysTrpAlaValSerSerHis 350
1001 AAGATCTATCTGGACAGACGCGCAGAGATATGCTTCTTACATCAT 1050
351 HisValIleCysGlnLeuLeuSerAspTrpLysGlnLysGlnMetLeu 367
1051 CATGTATTTGGCCAGTACTTCTTGACTCAAAAGAAACAGATGCTAA 1100
367 sLleSerSerGlyAsnSerAspProGluAsnValSerAArgThrArgAs 384
1101 AATCTCTTCGAAACAGCATTCAGAAATGTCTCAGAACCCAAATA 1150
384 ys 384
1151 AA 1152
seq_name: /SIDSI/gcgdata/hoid-geneseq/geneseq-emb1/NA2001A.DAT:AA167211
seq_documentation_block:
ID AA167211 standard; cDNA; 1155 BP.
XX
AC AA167211;
XX
DT 11-FEB-2002 (first entry)
XX
DE B305D isoform C splice variant 1 encoding cDNA.
XX
KW Genetic subtraction; DNA microarray analysis; polymerase chain reaction;
KW cancer; B305D; ss.
XX
OS Homo sapiens.
XX
FH Key
FT 1..1155 Location/Qualifiers
FT CDS /tag= a
FT /product= "B305D isoform C splice variant"
XX
PN MO200175171-A2.
XX
PD 11-OCT-2001.
XX
XX 02-APR-2001; 2001WO-US10631.
XX
XX 03-APR-2000; 2000US-194241P.
XX 20-JUL-2000; 2000US-219862P.
XX 27-JUL-2000; 2000US-221300P.
XX 18-DEC-2000; 2000US-256592P.
XX
PA (CORI-) CORIXA CORP.

```

XX Houghton RL, Dillon DC, Molesh DA, Xu J, Zehentner B, Persing DH;  
PI WPI; 2001-626449/72.  
XX P-PSDB; AAG65976.  
XX  
XX Identifying tissue (tumour)-specific polynucleotides overexpressed in  
PI tissue of interest as compared to control tissue, for detecting cancer  
PI cells in patient, comprises DNA microarray analysis or quantitative  
PI polymerase chain reaction -

XX Claim 4; Page 93-94; 127pp; English.

XX  
XX The invention relates to identifying tissue-specific polynucleotides (P)  
CC that involves performing a genetic subtraction to identify pool of (P)  
CC from tissue of interest (TI), performing DNA microarray analysis to  
CC identify first subset of polynucleotides (SP1) at least 2-fold over  
CC expressed in TI, and performing quantitative polymerase chain reaction  
CC (PCR) analysis on SP1 to identify second subset of (P). The method is  
CC useful for determining the presence or absence of a cancer cell in a  
CC patient, monitoring the progression of cancer in a patient using a  
CC biological sample such as blood, serum, lymph nodes, bone marrow, sputum,  
CC urine or a tumour biopsy sample. The methods are useful for determining  
CC the presence or absence of or monitoring progression of prostate, breast,  
CC colon, ovary, lung, head and neck, lymphoma, leukemia, melanoma, liver,  
CC gastric, kidney, bladder, pancreatic or endometrial cancer. The present  
CC sequence represents a cDNA encoding a B305D isoform C splice variant.

XX Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

#### alignment\_scores:

Quality: 2064.00 Length: 384  
Ratio: 5.375 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

#### alignment\_block:

US-09-810-936-304 x AAI67211 ..

Align seg 1/1 to: AAI67211 from: 1 to: 1155

1 MetValValGluValAspSerMetProAlaIleSerSerValIlyLysPr 17  
1 ATGGTGGTGGAGTTGATTCATGCGCGCTCTTCTGTGAGAAAGCC 50  
17 oPhedGlyLeuArgSerLysMetGlyLysTrpCysCysArgCysPheProC 34  
51 ATTGGTCTCAGAGCAAGATGGCAGATGGTGTGCTGCCGTCTCCCT 100  
34 YsCysArgGluSerGlyLysSerAsnValGlyThrSerGlyAspHisAsp 50  
101 GCTGCGAGGAGAGCGCAAGACACTGGGCACTTCTGGAGACCGAC 150  
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTrpCysArgH 67  
151 GACTCTGTATGAGACACTCAGAGCAAGATGGCAAGTGTGCCGCA 200  
67 sCysPheProCysArgGlySerGlyLysSerAsnValGlyAlaSerG 84  
201 CTGCTTCCCTGCTGCAAGGAGAGTGGCAAGACAGTGGCGCTTCTG 250  
84 LysPheHisAspAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100  
251 GAGACACAGAGCACTGTATGAAAGACTCAGAAACAAGATGGGCAAG 300  
101 TrpCysCysHisCysPheProCysCysArgGlySerGlyLysSerLys 117  
301 TGGTGTCTGCCACTGCTTCCCTGTGCAAGGAGGAGCGGCAAGAGG 350  
117 LglValAlaTrpLysAspTrpAspAspSerAlaPheMetGluProArgTyrH 134  
351 GGGGCGCTGGGAGACTAGCATGACAGTGCCTTCATGAGACCCAGGTACC 400

134 tSValArgGlyGluAspLeuAspLysLeuHisArgAlaAlaTrpTrpGly 150  
401 ACGTCGTGGAGAGATCTGGACAAAGCTCCACAGAGCTGCTGGTGGGCT 450  
151 LysValProArgLysAspLeuIleValMetLeuArgAspPhePheValAs 167  
451 AAAGTCCCCGAAAGAGATCTCATCTGATGCTCAGAGGACACTGAGTGA 500  
167 nLysLysAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsn 184  
501 CAAGAAAGACCAAAAGAGAGCTGCTACATCTGGCGCTTGGCAATG 550  
184 LysAsnSerGluValValLysLeuLeuLeuAspArgCysGlnLeuAsn 200  
551 GGAATTCAGAAATGTAATAAAGCTGCTGGACAGACGATGTCACACTTAA 600  
201 ValLeuAspAsnLysLysArgThrAlaLeuIleLysAlaValAlaGlnCysG 217  
601 GTCTTGTACACAAAAGAGAGAGCTGTGATTAAGGCGGTATGATGCA 650  
217 nGluAspGluCysAlaLeuMetLeuGlnHisGlyThrAspProAsn 234  
651 GGAAGATGAATGTCGTTAATGTTGCTGGAAATGCACTGATCCAAATA 700  
234 LProAspArgLysTrpGlyAsnThrThrLeuHisGlyAlaIleTyrAsnGlu 250  
701 TTCACAGTAGATGGAATATACCATCTGCACTAGCGTATCTATATGAA 750  
251 AspLysLeuMetAlaLysAlaLeuLeuLeuTyrGlyAlaAspIleGluSe 267  
751 GATAAATTAATGGCCAAAGACACTGCTTATGATGGTGGATATGGAATC 800  
267 rLysAsnLysHisGlyLeuThrProLeuLeuGlnGlyValHisGlnGln 284  
801 AAAAACAAGCATGGCTCTACACCACTGTTACTGGTGTACTGAGCAAA 850  
284 YsGlnGlnValValLysPheLeuIleLysLysLysAlaAsnLeuAsnAla 300  
851 AACAGCAAGTCGTGAATTTTAAATCAAGAAAAAGCGAATTTTAAATGCA 900  
301 LeuAspArgTrpGlyArgPheAlaLeuIleLeuAlaValCysCysGlySe 317  
901 CTGATATGATATGAAAGAGACTGCTCATACCTTGCGTATGTGTGGATC 950  
317 rAlaSerIleValSerLeuLeuGlnGlnAsnIleAspValSerSerG 334  
951 AGCAAGTATAGTCAGCCTTCTACTTGAAGCAAAATATGATGTATCTTCTC 1000  
334 LAspLeuSerGlyGlnThrAlaArgLysTrpAlaValSerSerHisHis 350  
1001 AAGATCATATGGACAGAGCGGCAAGAGATGCTGTTCTGATCATCAT 1050  
351 HisValIleCysGlnLeuLeuSerAspTrpLysGlnLysGlnMetLeuLys 367  
1051 CATGTAAATTTGGCAGATTAATCTTGTGACTACAAAGAAAACAATGCTTAA 1100  
367 sIleSerSerGluAsnSerAsnProGluAsnValSerArgThrArgAsnL 384  
1101 AATCTCTTTCGAAACAGCAATCCAGAAATGTCACAAACACAGAAATA 1150  
384 Ys 384  
1151 AA 1152

seq\_name: /STD/SI/gcgdata/hold-geneseq/geneseg-emb1/NA2001A.DAT:AA563807

seq\_documentation\_block:

ID AA563807 standard; cDNA, 1155 BP.

XX

AC AA563807;

XX

DT 29-JAN-2002 (first entry)

XX



1001 AAGATCTATCTGGACAGCGCCAGAGATGCTGTTCTAGTCATCAT 1050  
351 HisVal11ecySGInleuSerAspTyrLysGluLysGlnMetLeuLys 367  
1051 CAGGTAAATTTGCCAGTACTTCTGACTACAAAGAAAACACATGCTAAA 1100  
367 sileSerSerGluAsnSerAsnProGluAsnValSerArgThrArgAsnL 384  
1101 AATCTCTCTGAAACAGCAATCCAGMAAATGCTCCAGAACCAAGAAATA 1150  
384 ys 384  
1151 AA 1152

seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAH93714

seq\_documentation\_block:

ID AAH93714 standard; cDNA; 1155 BP.

AAH93714:

04-OCT-2001 (first entry)

Human prostate-specific cDNA sequence B305D splice variant #8.

Human: prostate cancer; prostate-specific; diagnosis; vaccine;

cytostatic; gene therapy; metastasis; ss.

Homo sapiens.

MO200151633-A2.

19-JUL-2001.

16-JAN-2001: 2001WO-US01574.

14-JAN-2000: 2000US-0483672.

(COR1) COR1XA CORP.

Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;

Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skelky YAW;

Wang A, Meagher MJ;

WPI; 2001-425873/45.

New polynucleotide encoding a prostate-specific protein, for

diagnosis, monitoring and treating prostate cancer in a patient and

for use in vaccines -

Claim 1, Page 347; 543pp; English.

The present invention describes polynucleotide sequences (I) which encode  
prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
and can be used in vaccine production and gene therapy. (I), (II),  
antibodies to (II), fusion proteins comprising (II), and isolated  
T cells prepared using (I) or (II) are used to treat cancer in a patient.  
(I) and the antibodies are also used in the detection of cancer in a  
patient. The cancer that is diagnosed or treated is particularly  
prostate cancer. (I) and (II) can be used in vaccines. The antibodies or  
(I) can be used for monitoring the progression of cancer in a patient.  
(I) and (II) can also be used to improve diagnostic and therapeutic  
methods for prostate cancer. They can indicate the level of metastasis  
as well as the prostate volume. AAH93714 to AAH93944 and AAH01115 to  
CC AAH01318 represent polynucleotide and amino acid sequences used in the  
exemplification of the present invention.

Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

alignment\_scores:

Quality: 2064.00

Ratio: 5.375

Length: 384

Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-810-936-304 x AAH93714 ..

Align seg 1/1 to: AAH93714 from: 1 to: 1155

1 MetValValGluValAlaSerMetProAlaIleSerValLysLysPr 17  
1 ATGGTGTTGAGGTTGATTCATGCCGGCTGCCCTCTTCGTGAAGAACCC 50  
17 ophEGlyLeuArgSerLysMetGlyLysTrpCysArgCysPheProC 34  
51 ATTTGGTCTCAGAGCAAGATGGGCAAGGTGTCGCCGTTCCTCCCT 100  
34 yscCysArgGluSerGlyLysSerAsnValGlyThrSerGlyAspHisAsp 50  
101 GCTGCAGGAGAGCGGCAAGACGATGGGCACCTTCGGAGACCAGAC 150  
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTrpCysArgHi 67  
151 GACTCTGCTATGAAGACACTCAGAGCAAGATGGGCAAGTGGTCCGCCA 200  
67 scYsPheProCysCysArgGlySerGlyLysSerAsnValGlyAlaSerG 84  
201 CTGCTTCCTCCCTGCTCAGGGGAGTGGCAAGACACGTCGGCGCTTCTG 250  
84 LysPheHisAspAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100  
251 GAGACCAAGACGACACTCTGCTATGAAGACACTCAGGACACAGATGGGCAAG 300  
101 TrpCysCysHisCysPheProCysArgGlySerGlyLysSerLysVal 117  
301 TGGTGTGCTGCCACTGCTTCCCTGCTGCTGAGGGAGCGGCAAGACGAGT 350  
117 LglValATrpgLysPTrAspPAspSerAlaPheMetGluProArgTyrH 134  
351 GGGCGCTTGGGAGACTACGATGACAGTGCCTTCATGAGCGCCAGGTAC 400  
134 lSValArgGlyGluAspLysPLeuAspLysLeuHisArgAlaAlaTrpTrpgLys 150  
401 ACGTCCGTGGAGAGATCTGGACAGAGCTCCACAGAGCTCCGTGGGGGT 450  
151 LysValProArgLysAspLeuIleValMetLeuArgAspTrAspValAs 167  
451 AAAGTCCCGCAGAAAGATCTCATCTGTCAGGCGACACTGAGCTGAA 500  
167 nLysLysAspLysGlnLysArgThrAlaLeuHisLeuHisLeuAlaSerAlaSng 184  
501 CAGAGAGACACAAAGAGAGCTGCTACATCTGGCTGGCTGGCCATG 550  
184 LysSerGluValValLysLeuLeuLeuAspArgCysGlnLeuAsn 200  
551 GGAATTCAGAAAGTAAACCTCTGCTGACAGACAGATGTCACACTATAT 600  
201 ValLeuAspAsnLysLysArgThrAlaLeuIleLysAlaValGlnCysGln 217  
601 GTCCTTGACACAAAGAGAGAGCTGATGAAGGCCGTTCATATGGCA 650  
217 nGluAspGluCysAlaLeuMetLeuGlnHisGlyThrAspProAsnI 234  
651 GGAAGATGAATGTGCGTTAATGTTGCTGGACATGAGCATGATCCAAATA 700  
234 lProAspGluTrpGlyLysAsnThrThrLeuHisThrAlaIleLysAsnGlu 250  
701 TTCAGATGAGTGTGAATACCACTGCTGACCTACGCTATCATATATAA 750  
251 AspLysLeuMetAlaLysAlaLeuLeuLeuTyrGlyAlaAspIleGluSe 267  
751 GATAAATTAATGSCAAAGCACGTGCTTATATGAGTGCATATGCAATC 800  
267 rLysAsnLysHisGlyLeuThrProLeuLeuLeuLysValHisGlnGlnL 284

801 AAAAACAAGCATGGCCATCAACACACTTACTTGGTGTACATGACAAA 850  
284 ysginglnvalvalyspheleullelyslsvalaspleuasnala 300  
851 AACGCAAGTCGTAATTTTAACTCAAAACCAATTTTAAATGCA 900  
301 leuaspargtyrlyargthralaleullealeuAvalcyscysglyse 317  
901 CTGATATGATATGAAAGACAGTGTCTCATCTGCTGTATGTGTGATC 950  
317 talaserllevalserleulleuulglnasnilleaspyalserSerg 334  
951 AGCAAGATAGTCAGCCCTTACTTACGCAAAATATGATGATCTCTC 1000  
334 lnaaspleuserglylnthrAlaarglutyralavalserSerhishts 350  
1001 AAGATCTATCTGGACAGACGCCAGAGATATGCTTCTTCAATCAT 1050  
351 HisvalillecysglnleuSerAspyrlyslgilylsglnmetleully 367  
1051 CATGTAAATTTGCCAGTTACTTCTGTACTCAAAAGAAAAACAGATGCTAAA 1100  
367 slleserSergluasnSerasnProgluasnValserArgThrargasnL 384  
1101 AATCTCTTGAACAGCAATGCAGAAATGTCTCAAGAACACAGAAATA 1150  
384 ys 384  
1151 AA 1152

seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAH85028  
seq\_documentation\_block:  
ID AAH85028 standard; cDNA; 1155 BP.  
XX AAH85028;  
AC  
XX  
XX  
DT 25-SEP-2001 (first entry)  
XX  
XX Human prostate-specific cDNA sequence B305D splice variant #8.  
DE  
XX Human prostate cancer; therapy: diagnosis; cat eye syndrome;  
KM chromosome 22q11.2; prostate-specific protein; chromosome 1;  
KM prostate specific antigen; PSA; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WC200134802-A2.  
PN  
XX 17-MAY-2001.  
PD  
XX 09-NOV-2000; 2000WO-US30904.  
PE  
XX 12-NOV-1999; 990S-0439313.  
PR 18-NOV-1999; 990S-0443686.  
PR  
XX (COR1-) COR1XA CORP.  
PA  
XX  
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
PI Kapos MD, Retter MW, Stolk JA, Day CH, Skeiky YMW, Wang A;  
XX WPI: 2001-308785/32.  
DR  
XX  
XX Isolated polypeptide comprising at least an immunogenic portion of a  
PT prostate-specific protein, useful in the diagnosis and therapy of  
PT prostate cancer -  
XX  
XX  
PS Claim 31: Page 246-247; 325pp; English.

CC The polypeptides, nucleic acids and antibodies from the present  
CC invention are useful in the diagnosis and therapy of prostate cancer.  
CC Prostate specific genes P704P, P712P, P774P, P775P and B305D are located  
CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome  
CC region. Prostate specific antigen (PSA) P501S was located on  
CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent  
CC polynucleotide and polypeptide sequences used in the exemplification  
CC of the present invention.  
XX  
SQ Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

Alignment\_scores:  
Quality: 2064.00 Length: 384  
Ratio: 5.375 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-810-936-304 x AAH85028 ..  
Align seg 1/1 to: AAH85028 from: 1 to: 1155

1 MetValValGluValAspSerMetProAlaAlaSerSerValIysLysPr 17  
|||||  
1 ATGGTGGTTGAGGTGATTCATGCGCGGCTCTTCTGTGAAGAAAGCC 50  
17 OPheGlyLeuArgSerIysMetGlyIysTrpCysArgCysPheProC 34  
51 APTTGGTCTCAGAGCAAGATGGCAAGTGGTGGCTGGCTTCCCT 100  
34 yscysArvgGluSerGlyLysSerAsnValGlyThrSerGlyAspHisAsp 50  
101 GCTGACAGGAGAGCGGCAAGACAGCAGTGGCATTCTGGAGACCAACAC 150  
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTrpCysArgHi 67  
151 GACTGTGCTATGAAGACACTGACGACCAAGATGGCAAGTGGTGGCCGA 200  
67 scysPheProCysCysArgGlySerGlyLysSerAsnValGlyAlaSerG 84  
201 CTGCTCCCTGCTGTCAGGGGAGTGGCAAGACAGCGGCGCTTCTG 250  
84 lYAspHisAspAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100  
251 GAGACACGACGACTGTGCTATGACACACTCAGGACAGATGGGCAAG 300  
101 TrpCysCysHisCysPheProCysCysArgGlySerGlyLysSerLysVa 117  
301 TGGTGTCTCCACTGCTTCCCTGCTGACAGGGGAGGCGCAAGCAAGGT 350  
117 lGlyAlaThrGlyAspTrpAspAspSerAlaPheMetGluProArgGlyRH 134  
351 GGGCGCTGGGGAGACTACGATGACATGCTCTCATGAGGCCAGAGTACC 400  
134 lSvalArgGlyGluAspLeuAspLysLeuHisArgAlaAlaIatPrpGly 150  
401 ACGTCCGTGGAGAGATCTGACACACTCCACGACACTCCTGCTGGGGT 450  
151 lYsValProArgLysAspLeulleValMetLeuArgAspThrAspValAs 167  
451 AAGTCCCGCAAGAGAGATCTCATGCTGATGCTGAGGAGACAGAGCTAAA 500  
167 nLysLysAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsng 184  
501 CAAGAGAGCAACCAAAAGAGACTGCTTACTTACCTGCGCTCTGCCAATG 550  
184 lYAsnSerGluValValLysLeulleuLeuAspArgArgGlyGlnLeuSn 200  
551 GGAATTCAGAGTGTAAAACTCTCTGTGACAGAGAGATGTAACCTTAAT 600  
201 ValLeuAspAsnLysLysArgThrAlaLeulleLysAlaValAlGlnCysG1 217  
601 GTCCTTGACACAAAGAGAGAGAGCTGTGATTAAGGCCGTAACTATGCCA 650

```

217 ngluAspGluCysAlaLeuMetLeuGluHisGlyThrAspProAsnI 234
|||||
651 GGAAGATGAATGTCGCTTAATGTTGCTGGACACGTGCACTGATCCAAAAA 700
|||||
234 leProAspGluTyrGlyAsnThrThrLeuHisTyrAlaIleTyrAsnGlu 250
|||||
701 TTCACAGTGAATGGAATAATACACCTCTGACACGCTATCATATATAGA 750
|||||
251 AspLysLeuMetAlaLysAlaLeuLeuTyrGlyAlaAspLysLeu 267
|||||
751 GATTAATTAATGATGCAAGACATGCTCTTATATAGGCGCATATGCAATC 800
|||||
267 TLYAsnLysHisGlyLeuThrProLeuLeuGlyValHisGluGlnL 284
|||||
801 AAAAAAACACATGGCTCTCACACCACTGTTACTTGSTGTACATGAGCAA 850
|||||
284 YSGInGlnValValLysPheLeuIleLysLysLysAlaAsnLeuAsnAla 300
|||||
851 AACAGCAAGTCGTCAAAATTTTATCAAGAAAAACGGAATTTAAATGCA 900
|||||
301 LeuAspArgTyrGlyArgThrAlaLeuIleLeuAlaValCysCysGlySe 317
|||||
901 CTGATATGATATGGAAGAGACTGCTCATACCTTGCTGTATGTTGGATC 950
|||||
317 AlaSerIleValSerLeuLeuGluGlnAsnIleAspValSerSerG 334
|||||
951 AGCAAGATATGTCAGCCTTCTACTTACGCAAAATATGATGATCTTCTC 1000
|||||
334 LAspLeuSerGlyGlnThrAlaArgGluTyrAlaValSerSerHisHis 350
|||||
1001 AAGATCTATCTGGACAGACGGCCAGAGATGCTGTTCTTCTAGTCATCAT 1050
|||||
351 HisValIleCysGlnLeuLeuSerAspTyrLysGluLysGluMetLeu 367
|||||
1051 CATTAATTTGGCCAGTTACTTCTGACTACAAAGAAAAACAGATGCTATA 1100
|||||
367 sIleSerSerGluAsnSerAsnProGluAsnValSerArgThrArgAsnL 384
|||||
1101 AATCTCTTCTGAAAAACAGCAATCCAGAAATGTCTCAGACACAGAAATA 1150
|||||
384 ys 384
||
1151 AA 1152

```

seq\_name: /sIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAH02779

seq\_documentation\_block:  
ID AAH02779 standard: cDNA: 1155 BP.  
XX  
AC AAH02779;  
XX  
DT 14-JUN-2001 (first entry)  
XX  
DE Prostate tumour antigen determined cDNA splice variant of B305D #B.  
XX  
KM Human: prostate tumour antigen: prostate tumour; therapy: diagnosis;  
KM prostate cancer; immunogenic; cytostatic; vaccine; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200125272-A2.  
XX  
PD 12-APR-2001.  
XX  
PF 04-OCT-2000; 2000WO-US27464.  
XX  
PR 04-OCT-1999; 99US-0157455.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Xu J, Skeiky YAM, Reed SG, Cheever MA.  
XX

```

DR MP1: 2001-245062/25.
DR P-PSDB; AAB74815.
XX
PT Prostate specific protein and its encoding polynucleotide, useful for
PT the treatment and diagnosis of prostate cancer -
PS
XX Claim 50; Page 231-232; 276pp; English.
XX
CC The present invention describes an isolated polypeptide (I) comprising
CC at least an immunogenic portion of a prostate tumour antigen protein or
CC its variant. (I) have cytostatic activity and can be used in vaccine
CC production. (I), prostate tumour antigen polynucleotides, an antigen
CC presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
CC pharmaceutical composition containing (I) are useful for inhibiting the
CC development of cancer in a patient. Antibodies specific for prostate
CC specific proteins and oligonucleotides that hybridize to a
CC polynucleotide that encodes a prostate specific protein are useful
CC for detecting the presence or absence of a cancer or monitoring the
CC progression of a cancer, especially prostate cancer.
CC AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences
CC used in the exemplification of the present invention.
XX
SQ Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;

```

alignment\_scores:  
Quality: 2064.00 Length: 384  
Ratio: 5.375 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-810-936-304 x AAH02779 ..

Align seg 1/1 to: AAH02779 from: 1 to: 1155

```

1 MetValValGluValAspSerMetProAlaAspSerValLysLysPr 17
|||||
1 ATGGTGCTTCAGGTTGATTCACATGCGCGCTGCTTCTGTCMAAACGC 50
|||||
17 oPhagLysLeuArgSerLysMetGlyLysTrpCysArgCysPheProC 34
|||||
51 ATTGTGCTCAGAGCAGACATGGCGCAAGTGCCTGCTTCCCTCCCT 100
|||||
34 yScysArgGluSerGlyLysSerAsnValGlyThrSerGlyAspHisasp 50
|||||
101 GCTGCAAGGAGAGCGGAGGAGCAAGCAAGCTGCGCATCTCGAACCACAC 150
|||||
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTrpCysArgHis 67
|||||
151 GACTCTGCTATGAAACACTCAAGACCAAGATGGCCAGTGGCGGCA 200
|||||
67 sCysPheProCysCysArgGlySerGlyLysSerAsnValGlyAlaSerG 84
|||||
201 CTGCTCTCCCTGCTGCGAGGGGAGTGGCAAGCAAGTGGCGCTCTCG 250
|||||
84 LysAsnHisAspAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100
|||||
251 GAGACACAGACGACCTGCTATGAAAGACACTCAGAACACAAATGGGCAAG 300
|||||
101 TrpCysCysHisCysPheProCysCysArgGlySerGlyLysSerLysVa 117
|||||
301 TGGTGGCTGCACATGCTTCCCTGCTGCAAGGGAGCGGCAAGCAAGGT 350
|||||
117 IGLValAlaTrpGlyAspTrpAspAspSerAlaPheMetGluProArgLys 134
|||||
351 GGGGGCTTGGGAGACTACGATGACAGTGGCTTCCAGGACCCAGAGTACC 400
|||||
134 IValArgGlyGluAspLeuAspLysLeuHisArgAlaAlaTrpTrpLys 150
|||||
401 ACGTCCGTGGAGAAAGATCTGACAAAGCTCCACAGAGATGCTGCGGGGT 450
|||||
151 LysValProArgLysAspLeuIleValMetLeuAlaGAspThrAspValAs 167
|||||

```



```

451 AAAGTCCCGAAGAAAGATCTCATGCTCATGCTCAGGAGACTGACGTGAA 500
167 nLysLysAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsnG 184
|||||
501 CAAAGAGACAGCAAGCAAAAGAGAGACTGCTCATGCTGAGCTCTGCAATG 550
184 LysSerGlnValValLysLeuLeuLeuAspArgArgCysGlnLeuAsn 200
|||||
551 GGAATTCAGAAAGTAGTAAATCTCTGCTGACAGCAGATGTCACCTTAAT 600
201 ValLeuAspAsnLysLysArgThrAlaLeuHisLeuAlaValAlaCysG 217
|||||
601 GTCCCTGACAAACAAAAGAGAGACGCTCGATAAAGCGCGTACAATGCA 650
217 nG1AspG1ucysAlaLeuMetLeuGlnHisGlyThrAspProAsn 234
|||||
651 GGAGAGATGATGTCGTTATGTTGCTGGACATGCGACTGATCCAAATA 700
234 LepProAspGluTyrGlyAsnThrThrLeuHisTyrAlaLeuTyrAsnG 250
|||||
701 TTCAGATGAGTATGGAATATCACCTCTGCTACGCTATCATATATATA 750
251 AspLysLeuMetAlaLysAlaLeuLeuLeuTyrGlyAlaAspIleGluSe 267
|||||
751 GATAAATTAAATGGCCAAAGACACTGCTTATATAGTGTGATATGAAATC 800
267 rLysAsnLysHisGlyLeuThrProLeuLeuGlnValHisGlnGln 284
|||||
801 AAAAACAAGCATGCGCTCACACCACTGTTACTGGTGACTGTGACAA 850
284 YsgInGlnValValLysPheLeuIleLysLysAlaAsnLeuAsnAla 300
|||||
851 AACAGCAAGTCGTGAATTTTATATCAAGAAAAAGCGAATTTAAATCA 900
301 LeuAspArgTyrGlyArgThrAlaLeuIleLeuAlaValCysCysGlySe 317
|||||
901 CTGATATATATGAGAGAGACTGCTCTCATACTTGGTATGTTGGATTC 950
317 rAlaSerIleValSerLeuLeuLeuGlnGlnAsnIleAspValSerSerg 334
|||||
951 AGCAAGATATAGTACGCCCTTCTACTTACGCAAAATATGATGATCTCTC 1000
334 LAspLeuSerGlnGlnThrAlaArgGluTyrAlaValSerSerHisHis 350
|||||
1001 AAGATCTATCTGACACAGCGCCAGAGATGCTGTTCTTCTGATCATCAT 1050
351 HisValIleCysGlnLeuLeuSerAspTyrLysGlnLysGlnMetLeuL 367
|||||
1051 CATGTAATTTGCCAGTTACTTCTGACTACAAAGAAAAACAGATGCTAAA 1100
367 sIleSerSerGlnAsnSerAsnProGlnAsnValSerArgThrArgAsn 384
|||||
1101 AATGCTCTCTGAAGACAGCAATCCAGAAATGTCTCAAGAACCAAAATA 1150
384 ys 384
1151 AA 1152
seq_name: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT: AAS9857
seq_documentation_block:
ID AAS9857 standard; cDNA; 1155 BP.
XX AAS9857;
AC
XX
XX 12-MAR-2002 (first entry)
XX Breast tumour-specific DNA B1Aq1 splice variant B1IC-15.
XX Human: breast cancer; PCR primer; ss; cytosstatic; immunostimulant;
XX tumour; vaccine; immunogenic.
OS Homo sapiens.

```

```

XX
PN WO200190152-A2.
XX
XX 29-NOV-2001.
PD
XX
PF 22-MAY-2001; 2001WO-US16776.
XX
PR 24-MAY-2000; 2000US-0577505.
PR 08-JUN-2000; 2000US-0590583.
PR 26-OCT-2000; 2000US-0699293.
PR 16-MAR-2001; 2001US-0810936.
XX
XX (CORI-) CORIXA CORP.
PA
PI Frudakis TN, Reed SG, Smith JM, Misher LE, Dillon DC, Retter MW;
PI Wang A, Skeiky YAM, Harlocker SL, Day CH;
XX
XX WPI: 2002-089919/12.
DR P-PDB; AAU74377.
DR
XX New breast tumour proteins and polynucleotides encoding them, useful for
PT treating and/or preventing cancer, particularly breast cancer, and for
PT eliciting humoral and/or cellular immune response
XX
PS Claim 1; Page 223; 245pp; English.
XX
CC The invention relates to novel breast tumour polynucleotides and
CC polypeptides. The polypeptides and polynucleotides are useful in
CC pharmaceutical compositions for treating and/or preventing cancer,
CC particularly breast cancer, and for eliciting an immune response,
CC may be used as probes or primers for nucleic acid hybridisation, in the
CC design and preparation of ribozyme molecules for inhibiting expression of
CC tumour polypeptides and proteins, and in recombinant DNA molecules to
CC direct expression of a polypeptide in host cells. AAS9570-AAS9888
CC represent novel human breast cancer protein coding sequences and
CC PCR primers of the invention.
XX
SQ Sequence 1155 BP; 346 A; 253 C; 297 G; 259 T; 0 other;
XX
alignment_scores:
Quality: 2064.00 Length: 384
Ratio: 5.375 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-810-936-304 x AAS9857 ..
Align seg 1/1 to: AAS9857 from: 1 to: 1155
1 MetValValGluValAspSerMetProAlaIleSerValLysLysPr 17
|||||
1 ATGGTGGTGAAGTGAATTCATGCGCGGCTGCTCTCTCTGTAAGAGCC 50
17 oPheGlyLeuArgSerLysMetGlyLysTyrCysArgCysPheProc 34
|||||
51 ATTGGTCTCAGAGCAAGATGGGCAAGTGGTGGCTGCTCTCCCT 100
34 YsCysArgGlnSerGlyLysSerAsnValGlyThrSerGlyAspHisAsp 50
|||||
101 GCTGAGAGAGAGCGGCAAGAGCAAGCTGGGCACTTCTGGAGACACAGC 150
101 GCTGAGAGAGAGCGGCAAGAGCAAGCTGGGCACTTCTGGAGACACAGC 150
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTyrCysArgH 67
|||||
151 GACTCTCTATGAAGACACTCAAGAGCAAGATGGGCAAGTGGTGGCGCA 200
67 sCysPheProCysCysArgGlySerGlyLysSerAsnValGlyAlaSerG 84
|||||
201 CTGCTTCCTGCTGCTGCAAGGAGAGTGGCAAGAGCAAGCTGGGCGTCTG 250
84 LysPheHisAspAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100
|||||

```

251 GAGACCACGACGACTCTGCTATGAAAGCACTCAGGAACAAGATGGCCAAAG 300  
101 TTPCYSCYSHISCYPheProCYSCysArgGlySerGlyLysSerIysVa 117  
301 TGGTGGTGGCACTGGTTCCCTGGCTGGCGGGGAGGCGGCAAGAGCAAGG 350  
117 TGTATATPGLYAspTYRAspAspSerAlaPheMetGluProArgTYR 134  
351 GGGCGGTGGGGAGACTAGCATGACAGTCCCTCATGGAGCCCAAGTACC 400  
134 TAVAlATrgLyLuAspLeuAspLysLeuHisArgAlaAlaIATrPYrGly 150  
401 AGGTGGGTGGAGACATCTGGACAAAGTCCACAGAGCTGGCTGGTGGGT 450  
151 LysValProArgLysAspLeuIleValMetLeuArgAspThrAspValAs 167  
451 AAAGTCCCGAAGAGATCTCTGCTCATGCTCAGGAGCACTGACGTGA 500  
167 nLysLysAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsn 184  
501 CAGAAGAGCAAGCAAAAGAGAGACTGCTCATCTGGCGCTGGCAATG 550  
184 LysAsnSerGluValValLysLeuLeuAspArgArgCysGlnLeuAsn 200  
551 GGAATTCAGAAAGTAAACTCTGCTGACAGACAGTGTCAACTTAAAT 600  
201 ValLeuAspAsnLysLysArgThrAlaLeuIleLysAlaValGlnCysG 217  
601 GTCTCTGACACAAAAGAGGACACTCTGTAAAGGCCGACAACTCCA 650  
217 nGluAspGluCysAlaLeuMetLeuGluHisGlyThrAspProAsn 234  
651 GGAAGATGAATGGCTTAACTTCTGCTGAAACATGGCACTATCCAATA 700  
234 LeProAspGluTYrGlyAsnThrLeuHisTYrAlaIleTYrAsnGlu 250  
701 TTCCAAATGATGTGAAATACCACTGCACTAGCAATCTATTAAGAA 750  
251 AspLysLeuMetAlaLysAlaLeuLeuTYrGlyAlaAspIleGlu 267  
751 GATTAATTAATGGCCAAAGCACTGCTTATATGCTGATATGCAATC 800  
267 TLYSAsnLysHisGlyLeuThrProLeuLeuGlnValHisGluGln 284  
801 AAAAACAACACATGGCCCTCACACACTGTACTTGGTATACATAGCAAA 850  
284 YSGlnGlnValLysPheLeuIleLysLysLysAlaAsnLeuAsnAla 300  
851 AACAGCAAGTCGTGAATTTTAAATCAAGAAAAAGCAATTTAAATGCA 900  
301 LeuAspArgTYrGlyArgThrAlaLeuIleLeuAlaValLysCysGly 317  
901 CTGGATGATATGGAAGACTGCTCTCATACTGCTGATATGTGGGATC 950  
317 TAlaSerIleValSerLeuLeuGluGlnAsnIleAspValSerSerg 334  
951 AGCAAGATATGTCAGCCTTCTACTTGAAGCAAAATATGATGATCTCTC 1000  
334 LnsAspLeuSerGlyGlnThrAlaArgGluTYrAlaValSerSerHis 350  
1001 AAGATCTATGTGACAGACGCGCAGAGATATGCTGTTCTAGTCATCAT 1050  
351 HisValIleCysGlnLeuLeuSerAspTYrLysGlnLysGlnMetLeu 367  
1051 CATCTATTTGGCAGTTACTTCTGACTACAAAGAAAAAGAGATGCTAAA 1100  
367 sLleSerSerGluAsnSerAspProLysValSerArgThrArgAsn 384  
1101 AATCTCTTGAAGACAGATCCAGAAATGTCTCAAGAACCAAGAAATA 1150  
384 YS 384  
1151 AA 1152

seq\_name: /SIDS1/gcdata/hold\_geneseg/geneseq-emb1/MA2002.DAT:AAS99869  
seq\_documentation\_block:  
ID AAS99869 standard; DNA: 1590 BP.  
XX  
AC AAS99869;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Breast tumour-specific gene B305D fusion construct.  
XX  
KW Human; breast cancer; PCR primer; ss; cytosolic; immunostimulant;  
KM tumour; vaccine; immunogenic.  
XX  
OS Homo sapiens.  
XX  
PN WO200190152-A2.  
XX  
PD 29-NOV-2001.  
XX  
PF 22-MAY-2001; 2001WO-US16776.  
XX  
PR 24-MAY-2009; 2000US-0577505.  
PR 08-JUN-2000; 2000US-0590583.  
PR 26-OCT-2000; 2000US-0699295.  
PR 16-MAR-2001; 2001US-0810936.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Fridakis TN, Reed SG, Smith JM, Misher LE, Dillon DC, Retter MW;  
PI Wang A, Skelky YAW, Harlocker SL, Day CH;  
XX  
DR WPI: 2002-089919/12.  
XX  
PT New breast tumour proteins and polynucleotides encoding them, useful for  
PT treating and/or preventing cancer, particularly breast cancer, and for  
PT eliciting humoral and/or cellular immune response  
XX  
PS Example 8: Page 235; 245pp; English.  
XX  
CC The invention relates to novel breast tumour polynucleotides and  
CC polypeptides. The polypeptides and polynucleotides are useful in  
CC pharmaceutical compositions for treating and/or preventing cancer,  
CC particularly breast cancer, and for eliciting an immune response,  
CC particularly humoral and/or cellular immune response. The polynucleotides  
CC may be used as probes or primers for nucleic acid hybridisation. In the  
CC design and preparation of ribozyme molecules for inhibiting expression of  
CC tumour polypeptides and proteins, and in recombinant DNA molecules to  
CC direct expression of a polypeptide in host cells. AAS99570-AAS99888  
CC represent novel human breast cancer protein coding sequences and  
CC PCR primers of the invention.  
XX  
SQ Sequence 1590 BP; 424 A; 403 C; 433 G; 329 T; 1 other;  
XX  
alignment\_scores:  
Quality: 2064.00 Length: 384  
Ratio: 5.375 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-810-936-304 x AAS99869 ..  
Align seg 1/1 to: AAS99869 from: 1 to: 1590  
1 MetValValGluValAspSerMetProAlaAlaSerSerValLysLysPr 17  
436 ATGGTGGTGAAGTGAATTCATGCGCGCTCTCTCTCTGTAAGAAACACC 485  
17 oPheGlyLeuArgSerLysMetGlyLysTrpCysCysArgCysPheProC 34  
486 ATTGTGTCAGAGCAAGATGGGCAAGTGGTGGCTGCTGCTTCCCTTCCCT 535

34 yscysaragluserglylyssersasnvalglythrserglyaspHisasp 50  
 536 GCTGCAGGAGGAGCGGCAAGACAGTGGCAGCTTCTGGAGACAGCAG 585  
 51 AspsersalmetlysrhrleuargserlysmetglylystrpCysaArgH 67  
 586 GACTGTGCTATGAGACACTCAGAGCAAGATGGCAAGTGTGTCCGCCCA 635  
 67 scyspHeprCysArgglySerglyLyssersasnvalglylaserg 84  
 636 CTGCTTCCCTGCTGCAGGGGAGGTGGCAAGCAAGTGGGCTCTCTG 685  
 84 LyspHisaspSersalmetlysrhrleuargasnlysmetglylys 100  
 686 GAGACCAAGACGACTGTCTATGAGACACTCAGAGCAAGATGGGCAAG 735  
 101 TrpCysArgHisCyspHeprCysArgglySerglyLysserlySva 117  
 736 TGGTGTGCTCAGCTCTCCCTGCTGCAGGGGAGCGGCAAGCAAGGT 785  
 117 LglYalatrpgLYaspTYaspSersalAphemetgluProargTYrH 134  
 786 GGGGGCTGGGGAGACTACGATGACAGYGCCCTTATGGAGCCAGGTACC 835  
 134 LsValatrgglygluaspLeuaspLysserHisArgAlaAlaTrpTrpGly 150  
 836 ACGTCCGTGGAGAAAGATCTGGACAAAGCTCCACAGAGCTGCTGGTGG 885  
 151 LysValProArgLysaspLeuileValmetLeuArgaspHrraspValas 167  
 886 AAAGTCCCGAGAAAGATCTCATGCTCATGCTCAGAGGACACTGACGTGAA 935  
 167 nLysLysaspLysgluLysArgThrAlaLeuHisLeuAlaSerAlaang 184  
 936 CAGAGAGGACAGCAAAAGAGGAGCTGCTACATCTGGCCTCTCCCATG 985  
 184 LysasnSerGluValValLysleuLeuLeuAspArgArgCysgluLeuasn 200  
 986 GGAATTCAGAAAGTAACTGCTGCGAGACAGCATGCTCAACTTAAT 1035  
 201 ValLeuaspasnLysLysArgThrAlaLeuileLysAlaValAGIncysgl 217  
 1036 GTCTTTCACACAAAAGAGAGAGCTCGATTAAGGCCGTACAAATGCCA 1085  
 217 ngluaspGluCysAlaLeuMetLeuLeuHisglYThrAspProasnI 234  
 1086 GGAAGATGAATGTGCTTAATGTCTGAGACATGGCAGTCAATA 1135  
 234 lerpAspArgLutyrGlyAsnThrThrLeuHisTYrAlaIleTYrAsnGlu 250  
 1136 TTCCAGATGAGTATGGAATACCACCTGTGCACCTACGCTATATATGAA 1185  
 251 AspLysleuMetAlaLysAlaLeuLeuLeuTYrGlyAlaaspIleGluSe 267  
 1186 GATTAATTAATGGCAAGCAGCTGCTTAATGTGCTGATATCGAATC 1235  
 267 rLYAsnLysHisglYleuThrProleuLeuLeuGluYAlaHisGluInL 284  
 1236 AAAAANAACATGAGCTGCACACACACTGTACTTGTGTATATAGCAAA 1285  
 284 ysgIngluValValLyspHeleuLleLysLysLysAlaasnLeuAsnAla 300  
 1286 AACGCAAGTCGTGAATTTTAAATCAAAAAAAGCAAAATTTTAAATGCA 1335  
 301 LeuaspArgTYrGlyArgThrAlaLeuileLeuAlaValCysCysGlySe 317  
 1336 CTGATATGATATGAGAGAGAGTGTCTCATACTTGTGTATGTGTGGATC 1385  
 317 rAlaSerlleValSerleuLeuLeuGluInasnIleAspValSerSerG 334  
 1386 AGCAAGATATGATGAGCTTCTACTTGAAGCAAAATATGATGTATCTCTC 1435

334 lnaaspLeuSerGlyGlnThrAlaArgglutYrAlaValSerSerHisHis 350  
 1436 AAGATCTATCTGACAGACAGCGCAGAGATGCTGTTTCTAGTCACAT 1485  
 351 HisValIleCysGlnleuLeuSerAspTYrLysgluLysgluMetLeu 367  
 1486 CAGTGTATTTGGCAGTTACTTCTGTACTACAAAGAAAAAGATGTAA 1535  
 367 sLleSerSerGluAsnSerAsnProGluAsnValSerArgThrArgAsnL 384  
 1536 AATCTCTTGTGAAGACGCAATTCAGAAAAATGCTCAAGAACAGAAATA 1585  
 384 Ys 384  
 1586 AA 1587  
 seq\_name: /stidsl/gcgdata/hold-geneseq/geneseqn-emb1/MA2002.DAT:AA599872  
 seq\_documentation\_block:  
 ID AA599872 standard; DNA; 1155 BP.  
 XX  
 AC AA599872;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Breast tumour-specific gene B305D homologue #2.  
 XX  
 KW Human; breast cancer; PCR primer; ss; cytostatic; immunostimulant;  
 KW tumour; vaccine; immunogenic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200190152-A2.  
 XX  
 PD 29-NOV-2001.  
 XX  
 PE 22-MAY-2001; 2001WO-US16776.  
 XX  
 PR 24-MAY-2000; 2000US-0577505.  
 XX  
 PR 08-JUN-2000; 2000US-0590583.  
 XX  
 PR 26-OCT-2000; 2000US-0699295.  
 XX  
 PR 16-MAR-2001; 2001US-0810936.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Frudakis TN, Reed SG, Smith JM, Misher LE, Dillon DC, Retter MW;  
 PI Wang A, Skeiky YAW, Harlocker SL, Day CH;  
 DR MPI; 2002-089919/12.  
 DR P-PSDB: AAU74390.  
 XX  
 PT New breast tumour proteins and polynucleotides encoding them, useful for  
 PT treating and/or preventing cancer, particularly breast cancer, and for  
 PT eliciting humoral and/or cellular immune response  
 XX  
 PS Claim 1; Page 239; 245pp; English.  
 XX  
 CC The invention relates to novel breast tumour polynucleotides and  
 CC polypeptides. The polypeptides and polynucleotides are useful in  
 CC pharmaceutical compositions for treating and/or preventing cancer,  
 CC particularly breast cancer, and for eliciting an immune response,  
 CC particularly humoral and/or cellular immune response. The polynucleotides  
 CC may be used as probes or primers for nucleic acid hybridisation, in the  
 CC design and preparation of ribozyme molecules for inhibiting expression of  
 CC tumour polypeptides and proteins, and in recombinant DNA molecules to  
 CC direct expression of a polypeptide in host cells. AA599570-AA599888  
 CC represent novel human breast cancer protein coding sequences and  
 CC PCR primers of the invention.  
 XX  
 SQ Sequence 1155 BP; 346 A; 253 C; 296 G; 260 T; 0 other;  
 alignment\_scores:

Quality: 2054.00 Length: 384  
Ratio: 5.349 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.479

alignment\_block:  
US-09-810-936-304 x AAS99872 ..

Align seg 1/1 to: AAS99872 from: 1 to: 1155

```

1 MetValValAlGluValAspSerMetProAlaAlaSerValIleLysLysP 17
1 ATGGTGGTTGAGGTGATTCATCCATGCCGGCTGCTCTCTCTGTGAAGAAAGCC 50
17 oPhaGlyLeuArcSerLysMetGlyLysTrpCysCysArgCysPheProc 34
1 ATTTGGTCTCAGAGCAAGATGGCAGATGGTGGTGGTGGTGGTGGTGGTGGTGGT 100
34 yscysArgGluSerGlyLysSerAsnValGlyThrSerGlyAspHisAsp 50
101 GCTGCAGGGAGAGCGGCAAGCAACGTGGCGACTCTGGAGACCAAGAC 150
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTrpCysArgH 67
151 GACTCTGCTATGAGAGCACTCAGAGCAAGATGGGCAAGTGGTGGCCGCA 200
67 scysPheProcCysArgGlySerGlyLysSerAsnValGlyLysArg 84
201 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 250
84 LysPheAspAspSerAlaMetLysThrLeuArgAsnLysMetGlyLys 100
251 GAGACCAAGCAAGCACTCTGCTATGAAGACACTCAGAACACATGATGGCAAG 300
101 TrpCysCysHisCysPheProcCysArgGlySerGlyLysSerLysVa 117
301 TGGTGGTGGCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 350
117 LglValAlaTrpGlyAspTrpAspAspSerAlaPheMetGluProArgTrpH 134
351 GGGGGCTGGGAGAGACTACGATGACAGTCCCTTCATGAGGCCAGGTACC 400
134 lsvAlaArgGlyLysLeuAspLysLeuHisArgAlaAlaAlaTrpTrpGly 150
401 ACGTCCCTGGAGAAAGATCTGGACAAGCTCCACAGAGCTGGCTGGGGGT 450
151 LysValProArgLysAspLeuIleValMetLeuArgAspTrpAspValas 167
451 AAGTCCCGCAAGAAAGATCTCATCGTCATGCTCAGGGACACTGACGTGAA 500
167 nllysLysAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsn 184
501 CAAGCAGACAGCAAGCAAGAGAGAGCTCTCTACATCTGGCTCTGCCAATG 550
184 LysAsnSerGluValAlaLysLeuLeuLeuAspArgCysGlnLeuAsn 200
551 GGAATTCAGAAAGTAAATAAATCTCTCTGAGACAGACGATGCAACTTAAT 600
201 ValLeuAspAsnLysLysArgThrAlaLeuIleLysAlaValGlnCysG 217
601 GTCCTTGACAAACAAAAGAGAGAGCTCTGATTAAGGCCGATCAATGCA 650
217 ngluAspGluCysAlaLeuMetLeuGlnHisGlyThrAspProAsn 234
651 GGAAGATGATGTGCTTAATGTGTGTGACATGCACTGATCCAAATA 700
234 leproAspGluTrpGlyAsnThrThrLeuHisTrpAlaIleTrpAsnGlu 250
701 TTCAGATGAGTATGAAATACCACTCTCAGTACATATATATAATGAA 750
251 AspLysLeuMetAlaLysAlaLeuLeuLeuTrpGlyAlaAspIleGluSe 267
751 GATTAATTAATGGCAAGACACTGCTCTTATATGTGTGCTGATATGCAATC 800

```

```

267 rlyAsnLysHisGlyLeuThrProLeuLeuGlnValAlaHisGlnGlnL 284
801 AAAAACAAGCATGGCTCACAACACTGTACTTGTGTGTACATAGCAAA 850
284 ysgInGlnValValLysPheLeuIleLysLysAlaAsnLeuAsnAla 300
851 AACAGCAAGTCGTGAATAATTTAATTAAAGAAAAGCAAAATTTAAATGCA 900
301 LeuAspArgTrpGlyArgThrAlaLeuIleLeuAlaValCysCysGlySe 317
901 CTGGATAGATATGGAAGAGACTGCTCTCATACTTGGCTGTATGTGGATC 950
317 rAlaSerIleValSerLeuLeuGlnGlnAsnIleAspValSerSerg 334
951 AGCAAGTATAGTCAGCTTCTACTTACGCAAAATATTGATGATCTTCTC 1000
334 lnaAspLeuSerGlyGlnThrAlaArgGluTrpAlaValSerSerHis 350
1001 AAGATCTATCTGACACAGCGCCAGAGATATGCTGTTAGTATCATAT 1050
351 HisValIleCysGlnLeuLeuSerAspTrpLysGluLysGlnMetLeu 367
1051 CATCTAATTTGCCAGTACTTCTGTACTCAAAAGAAAAGATGCTTAAA 1100
367 sIleSerSerGluAsnSerAsnProGluAsnValSerArgTrpArgAsn 384
1101 AATCTCTTCTGAAACACGATCAGAAATGCTTCAGAACCAAGAAATA 1150
384 ys 384
1151 AA 1152
seq_name: /SDSL/gcgdata/hold-geneseg/genesegn-emb1/NA2000.DAT.AAC81012
seq_documentation_block:
ID AAC81012 standard; CDNA; 2000 BP.
XX
AC AAC81012;
XX
DT 13-FEB-2001 (first entry)
XX
DE Human B1Agl1 antigen splice isoform B1A1C-8 cDNA.
XX
KW Human; breast tumour-specific antigen; cytostatic; vaccine;
KW breast cancer; B1Agl1; B1Agl1; B15Agl1; ss.
XX
OS Homo sapiens.
XX
PN W0200061753-A2.
XX
PD 19-OCT-2000.
XX
PF 07-APR-2000; 2000MO-0509312.
XX
PR 09-APR-1999; 99US-0289198.
PR 28-OCT-1999; 99US-0429755.
PR 23-MAR-2000; 2000US-0534825.
XX
PA (COR1-) COR1XA CORP.
XX
PI Frudakis TN, Smith JM, Reed SG, Misher LE, Retter MM, Dillon DC;
XX
WP1; 2000-628403/60.
XX
DR P-PSDB; AAB28629.
XX
PT An isolated polypeptide comprising an immunogenic portion of a breast
PT tumor protein used for inhibiting the development of cancer, especially
PT breast cancer, and monitoring cancer progression in a patient -
XX
PS Claim 4; Page 177-178; 187pp; English.
XX
CC The present sequence is given in a specification relating to compositions
CC and methods for the treatment and diagnosis of breast cancer. Nucleotide

```



PR 27-JUL-2000; 2000US-221300P.  
PR 18-DEC-2000; 2000US-256592P.

PR 18-DEC-2000; 2000US-256592P.

PA (CORI-) CORIXA CORP.

PI Houghton RL, Dillon DC, Molesh DA, Xu J, Zehentner B, Persing DH;

DR WPI: 2001-626449/72.

DR P-PSDB; AAG65977.

PT Identifying tissue (tumour)-specific polynucleotides overexpressed in PT tissue of interest as compared to control tissue, for detecting cancer cells in patient, comprises DNA microarray analysis or quantitative polymerase chain reaction -

PS Claim 4; Page 94; 127pp; English.

CC The invention relates to identifying tissue-specific polynucleotides (P)  
CC that involves performing a genetic subtraction to identify pool of (P)  
CC from tissue of interest (TI), performing DNA microarray analysis to  
CC identify first subset of polynucleotides (SP1) at least 2-fold over  
CC expressed in TI, and performing quantitative polymerase chain reaction  
CC (PCR) analysis on SP1 to identify second subset of (P). The method is  
CC useful for determining the presence or absence of a cancer cell in a  
CC patient, monitoring the progression of cancer in a patient using a  
CC biological sample such as blood, serum, lymph nodes, bone marrow, sputum  
CC urine or a tumour biopsy sample. The methods are useful for determining  
CC the presence or absence of or monitoring progression of prostate, breast  
CC colon, ovarian, lung, head and neck, lymphoma, leukemia, melanoma, liver  
CC gastric, kidney, bladder, pancreatic or endometrial cancer. The present  
CC sequence represents a cDNA encoding a B305D isoform C splice variant.

Sequence 2000 BP; 698 A; 388 C; 489 G; 425 T; 0 other;

alignment_scores:	
Quality:	2024.00
Ratio:	5.383
Percent Similarity:	100.000
Gaps:	0
Percent Identity:	100.000

alignment\_block:

Align seg 1/1 to: AA167212 from: 1 to: 2000

[illegible]

117	1GLVATRPGLSPYRPAAPSPERALAPhmeCiluprOargryrH	134
351	GGGCCCTTGGGAGACTACGATGACAGGCGCTTATGAGGCCACGATACC	400
134	ISVALARGIYGLUASPLeuASPlyLeuHIIsARGALALATPTrPcLY	150
401	ACGTCCTGGAGAGATCTGGACAAAGCTCCACAGAGCGCTCGTGGGGT	450
151	LYSVALIProVALysASPLeuILEVALMetLeuHrGAsPTrPAsPVALIs	167
451	AAAGTCCCCAGAAAGATCTCATGTGTGTCAGAGGACACTGACGTGAA	500
167	nLYSASPlySGlnLYSArgTHRAlaLeuHISLeuAlaSerAlaAsnG	184
501	CAAGAGAGCAAGCAAAAGAGAGCTCTACATCTGGCGCTTCCCAATG	550
184	LYASnSergIValValLYSLeuLeuLeuAsPARGATCySGlnLeuAsn	200
551	GSAAATTCAGAGTAGTAACCTCGGCTGGACAGACGATGTCACATTAA	600
201	VALLeuAsPAsnLYSArgTHRAlaLeuILEuSAlaValGlnCYSGI	217
601	GGCTCTTGACAAACAAAAGAGGACAGCTGTGATAAGGCCGATACATGCCA	650
217	nGLUAsPGLUCYSAlaLeuMetLeuLEuGLAHISGlnTHrAsPProAsnI	234
651	GSAAATGATAAATCGCTTAATGTTCTGGACACATGGCACTGATCCAAATA	700
234	LeProASGlnTrpCYLysAsnTHRTrPLeuHISpTYRAlaILETYRAsnLU	250
701	TTCCAGATGAGAGATGGAAATACCACTCTGCACTACGATACATATATGAA	750
251	ASPlySLeuMetAlaLYSAlaLeuLeuLeuTYRGLYAlaAsPILeGLuSe	267
751	GATTAATTAATAGGCCAAAGCACCTGCTTTAATAGTGTGATATCGAATC	800
267	TrYSAsnLYSHISGlyLeuTHRProLeuLeuLeuGlnLYAlaHISGlnGlnI	284
801	AAAAACAACAGCATGGCCCTACACACCACTGTACTTGTTGGTACATGAGCAA	850
284	ysGlnGlnValLYSPhelLeuILEuLYSLYSAlaAsnLeuAsnAla	300
851	AAACACCAAGTCTGTAAATTTTAAATCAAGAAAAAGCCAAATTTAATATCCA	900
301	LeuAsPARGTYRGLYARGTHRAlaLeuLeuLeuAlaValCYSCysGLYSe	317
901	CTGGATAGATATGAGAGAGACTCTCTCATCTGCTGATGTGTGGAGATC	950
317	TrAlaSerILEValSerLeuLeuLeuGlnGlnAsnILEAsPValSerSerg	334
951	AGCAAGTATAGTCAGCCCTTCTACTTGAGCAAAATATTTGATGTACTTCTC	1000
334	IASPLeuSergLYClnTHRAlaArgGLYRAlaValSerSerHISHis	350
1001	AGAGTCATCTAGACAGACGGCCAAAGAGATAGCGTTTCAGATCATAT	1050
351	HISValILECYGlnLeuLeuSergAsPTrpLYSGlnLYSGlnMetLeuLY	367
1051	CATGTATATTGGCACTTACTTCTGTACACAAAGAAAAACAGATGCTAAA	1100
367	sILeSerSergLUAsnSergAsnProGLu	376
1101	AAATCTCTCTGAAAAACAGCAATCCAGAA	1128

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAS63808

```
seq_documentation_block:
```

ID AAS63808 standard; cDNA; 2000 BP.

XX

AC AAS63808;

XX  
XX

XX







851 AACAGCAAGCTCGTAATTTTAAATCAAGAAAAAGCAATTTAAATGCA 900  
 301 LeuAspArgTyrGlyArgThrAlaLeuIleuAlaValCysCysGly 317  
 901 CTGGATAGATGATGAAGGACCTGCTCTACTTCTGCTATGTTGTGGATC 950  
 317 rAlaSerIleValSerIleuLeuGluGlnAsnIleAspValSerSerG 334  
 951 AGCAGATATAGTCAGCCTTCTACTGAGCAAAATATGATGATCTTCTC 1000  
 334 lAspLeuSerGlyGlnThrAlaArgGluTyrAlaValSerSerHisHis 350  
 1001 AAGATCTACTGTGACAGCGCCACAGAGATGCTGTTCTAGCATCAT 1050  
 351 HisValIleCysGlnIleuLeuSerAspTyrIlysgIuIysGlnMetLeu 367  
 1051 CATGTATTTGGCAGTTACTTCTCTACTACAAAGAAAAACAGATGCTAAA 1100  
 367 sIleSerSerGluAsnSerAsnProGlu 376  
 1101 AATCTCTTCTGAAAACAGCAATCCAGAA 1128  
 seq\_name: /SIDSI/gcgdata/hold-geneseg/geneseqn-emb1/NA2001A.DAT:AAH85029  
 seq\_documentation\_block:  
 ID AAH85029 standard; cDNA; 2000 BP.  
 XX AAH85029;  
 AC  
 XX AAH85029;  
 DT 25-SEP-2001 (first entry)  
 XX  
 DE Human prostate-specific cDNA sequence B305D splice variant #9.  
 XX  
 KW Human; prostate cancer; therapy; diagnosis; cat eye syndrome;  
 KM chromosome 22q11.2; prostate-specific protein; chromosome 1;  
 XX prostate specific antigen; PSA; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200134802-A2.  
 XX  
 PD 17-MAY-2001.  
 XX  
 PF 09-NOV-2000; 2000WO-US30904.  
 XX  
 PR 12-NOV-1999; 99US-0439313.  
 PR 18-NOV-1999; 99US-0443686.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
 PI Kalos MD, Retter MW, Stolk JA, Day CH, Skelky YAW, Wang A;  
 XX WPI: 2001-308785/32.  
 DR  
 XX  
 XX Isolated polypeptide comprising at least an immunogenic portion of a  
 PT prostate-specific protein, useful in the diagnosis and therapy of a  
 PT prostate cancer -  
 XX  
 PS Claim 31: Page 247-248; 325pp; English.  
 XX  
 CC The present invention describes an isolated polypeptide (P1) comprising  
 CC at least an immunogenic portion of a prostate-specific protein, or its  
 CC variant. Also described are polynucleotides (N1) encoding (P1). (P1) and  
 CC (N1) have cytostatic activity and can be used in vaccine production.  
 CC The polypeptides, nucleic acids and antibodies from the present  
 CC invention are useful in the diagnosis and therapy of prostate cancer.  
 CC Prostate specific genes PT04P, P712P, P774P, P775P and B305P are located  
 CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome  
 CC region. Prostate specific antigen (PSA) P501S was located on  
 CC chromosome 1. AAH84671 to AAH85143 and AA69900 to AA69907 represent  
 CC polynucleotide and polypeptide sequences used in the exemplification  
 CC of the present invention.

XX  
 SQ Sequence 2000 BP; 698 A; 388 C; 489 G; 425 T; 0 other;  
 alignment\_scores:  
 Quality: 2024.00 Length: 376  
 Ratio: 5.383 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000  
 alignment\_block:  
 US-09-810-936-304 x AAH85029 ..  
 Align seg 1/1 to: AAH85029 from: 1 to: 2000  
 1 MetValValGluValAspSerMetProAlaIleSerSerValIlyAspPr 17  
 1 ATGGTGTTGAGGTTGATTTCATCCGCGCTCTCTTCTGTGTAAGAGCC 50  
 17 oPhcGlyLeuArgSerIlySerIlyMetGlyIlystrPcYsArGcYsPheProc 34  
 51 ATTTGGTCTCAGAGAGCAAGATGGGCAAGTGCTGCTCCGCTTCTCCCT 100  
 34 YsCysArgGluSerGlyIlySerSerAsnValGlyThrSerGlyAspHisAsp 50  
 101 GCTGCAGGGAGAGCGGCAAGACAGCTGGCGACTTCTGAGACACAGCAGC 150  
 51 AspSerAlaMetIlyThrLeuArgSerIlyMetGlyIlystrPcYsArgHi 67  
 151 GACTCTGCTATGAAACACTCAGGAGCAAGATGAGCAAGCTGATGCGGCA 200  
 67 sCysPheProcYsCysArgGlySerGlyIlySerSerAsnValGlyAlaSerG 84  
 201 CTGCTCCCTCTCTCAGGGAGGAGTGGSCAAGACACAGTGGCGCTTCG 250  
 84 IlyAspHisAspSerSerAlaMetIlyThrLeuArgSerIlyMetGlyIly 100  
 251 GAGACACAGCAGCTCTCTATGAGACACTCAGGACAGCAAGATGGGCAAG 300  
 101 TrpCysCysHisCysPheProcYsCysArgGlySerGlyIlySerIlyVal 117  
 301 TGGTGTGTCACACTGCTCCCTCTGTCAGGGAGGAGCGGCAAGCAAGGT 350  
 117 lGlyAlaTrpPcGlyAspTyrAspSerSerAlaPheMetGluProArgTyrH 134  
 351 GGGCGCTTGGGAGACTGACATGACACTGCTTCACTGAGGCCAGGTACC 400  
 134 lValAlaArgGlyGluAspLeuAspIlySerIlyHisArgAlaAlaTrpPcGly 150  
 401 ACGTCCGTGAGAAAGATCTGGCAAGCTCCACAGACACTGCTGGTGGGT 450  
 151 lYsValProArgIlyAspLeuIleValMetLeuArgAspThrAspValAs 167  
 451 AAGTCCCCAGAAAGATCTCATCTGCTCAGGAGACACAGTACAGTGAA 500  
 167 lYsIlyAspIlysgIuIlyAspArgThrAlaLeuHisIleuAlaSerAlaAsnG 184  
 501 CAAAGAAAGCAAGCAAAAGAGAGACTGCTCTACATCTGGCTCTGCCAATG 550  
 184 lYsAsnSerGluValValIlyLeuLeuLeuAspArgArgCysGlnIleuAsn 200  
 551 GGAATTCACAAAGTACTAAACTCTGCTGACAGAGAGAGTGAACCTTAAT 600  
 201 ValLeuAspAsnIlySArgThrAlaLeuIleValAlaValGlnCysGln 217  
 601 GTCCTTGACAAACAAAAGAGAGAGCTGATTAAGCCGTCATATGCCA 650  
 217 nGluAspGluCysAlaLeuMetLeuLeuGluHisGlyTyrAspProAsnI 234  
 651 GGAAGATGATGTGCGTTAAATGTGCTGGACATGGCAGATCAAAATA 700  
 234 lProAspGluTyrGlyIlyAsnThrThrLeuHisTyrAlaIleTyrAsnGlu 250  
 701 TTCAGATGAGTATGGAATACCACTCTGCACATACCTATCTAATATGAA 750

251 AspLysLeuMetAlaLysAlaLeuLeuTyrGlyAlaAspIleGluSe 267  
|||||  
751 GATAAATTAAAGCCAAAGCACTGCTTATATAGGCGTGAATATCGAATC 800  
|||||  
267 ILysAsnLysHisGlyLeuThrProLeuLeuGlyValHisGluGlnL 284  
|||||  
801 AAAAACAAGCATGGCCCTCACACCACTGTTACTGTGTGATGAGCAAA 850  
|||||  
284 YSGInGlnValValLysPheLeuIleLysLysLysAlaAsnLeuAsnAla 300  
|||||  
851 AACAGCAAGTGTGAAATTTTAAATCAAGAAAAAACCAATTTAAATGCA 900  
|||||  
301 LeuAspArgTyrGlyArgThrAlaLeuIleLeuAlaValCysCysGlySe 317  
|||||  
901 CTGGATAGATATGAGAGAGACTGCTCATACTTGCATATGTTGTGGATC 950  
|||||  
317 FAlaSerIleValSerLeuLeuGluGlnAsnIleAspValSerSerg 334  
|||||  
951 AGCAAGTATAGTCAGCCTTCTACTGTAGCAAAATATGATGATCTCTC 1000  
|||||  
334 LnaSpLeuSergGlyGlnThrAlaArgGluTyrAlaValSerSergHis 350  
|||||  
1001 AAGATCTATCTGGACAGACGGCCAGAGATGCTGTTCTAGTCATCAT 1050  
|||||  
351 HisValIleCysGlnLeuLeuSerAspTyrLysGluLysGlnMetLeuLy 367  
|||||  
1051 CATGTAATTTGCCAGTTACTTCTGACTACAAAGAAAAACAGATGCTTAA 1100  
|||||  
367 sIleSergGluAsnSergAsnProGlu 376  
|||||  
1101 AATCTCTCTGAAAAACAGCAATCCAGAA 1128



```

223 eumelleuleugluh1sglythrAspProasn1leProAspGluTyrGly 239
|||||
102 TAAATGTTGTGAACATGGACATGATCCAAACATTCGAGATGAGATGGA 151
|||||
240 AsphThrleuH1sTyrAla1leTyrAsnGluAspLysIleuMetAlaIy 256
|||||
152 AATACCACTGACATGACCTATCTCAATGAAATGAAATTAATGAGCAA 201
|||||
256 salaleuleuleutyrGlyAlaAsp1leGluSerLysAsnLysGlyL 273
|||||
202 AGCACTGCTCTATACGGTGTGATATCGAATCAAAAACAGCATGGCC 251
|||||
273 eutThrProleuleuGlyValH1sGluGlnLysGlnValIValLys 289
|||||
252 TCACACCACTGCTACTTGGTGTACAGACCAAAAACAGCAAGTGCGAAA 301
|||||
290 PheleuleuleysLysAlaAsnLeuAsnAlaleuAspArgTyrGlyAr 306
|||||
302 TTTTAAATCAAGAAAAAGCGAAATTTAAATGCGGTGATATATGGAAG 351
|||||
306 gThrAlaLeu1leleuAlaValCysCysGlySerAlaSer1leValSerL 323
|||||
352 AACTGCTCTCATPACTGCTGTATGTTGTGATCAGCAAGTATAGTCAACC 401
|||||
323 euleuleuGluGlnAsn1leAspValSerSerGlnAspLeuSerGlyGln 339
|||||
402 TTCTACTGAGCAAAATATGATGATCTTCTCAAGATCTATCTGACAG 451
|||||
340 ThrAlaArgGluTyrAlaValSerSerH1sH1sVal1leCysGlnLe 356
|||||
452 ACGGCCAGAGATGATGCTGTTTCTACTCATCATCATGTAAT.TGCCACT 500
|||||
356 uleuSerAspTyrLysGluLysGluMetleuleys1leSerSerGluAsn 373
|||||
501 ACTTTCGACTACAAAGAAAAACAGATGCTAAATCTCTCTGAAAAACA 550
|||||
373 erAsnProGluAsnValSerArgThrArgAsn 383
|||||
551 GCAATCCAGAAAATGTTCTCAAGAACAGAAAT 582
|||||
seq_name: gb_est2:BM469654

seq documentation block:
LOCUS BM469654 1020 bp mRNA linear EST 05-FEB-2002
DEFINITION AGEMCOURT_6444673 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5585453
5' mRNA sequence.
ACCESSION BM469654
VERSION BM469654.1 GI:18518696
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1020)
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNLT at:
http://image.llnl.gov
Plate: ILNLT351 row: m column: 06
High quality sequence stop: 680.
location/qualifiers
1..1020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5585453"

```

```

/clone_id="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC library."
BASE COUNT 348 a 208 c 218 g 240 t 6 others
ORIGIN

alignment_scores:
Quality: 597.00 Length: 257
Ratio: 3.411 Gaps: 4
Percent Similarity: 68.093 Percent Identity: 49.805

alignment_block:
US-09-810-936-304 x BM469654 ..

Align seq 1/1 to: BM469654 from: 1 to: 1020

134 HisValArgGlyGluAspLeuH1sArgValAlaIArgTyrGly 150
|||||
2 CACGTCCGAGACCGAGATCTCGGCAAGATCCACAAAGCTGCCAGCGGG 51
|||||
150 LysValProArgLysAspLeu1leValMetLeuArgAspThrAspVal 167
|||||
52 TAAATGTCGCAAAAGTCACAGATCTTTGTCTGAGAAAGATGGCTTGA 101
|||||
167 snLysAspLysGlnLysArgThrAlaLeuH1sLeuAlaSerAlaAsn 133
|||||
102 ACGATACAGCAAGATGACAGAGAGCGCTACATTTGGCTGGCCAAAT 151
|||||
184 GlyAsnSerGluValValLysLeuLeuAspArgGlyCysGlnLeuAs 200
|||||
152 GGTGATCCAGAAAGTAACTCTCTGTGGACAGAAAATGCCAGCTCAA 201
|||||
200 nValLeuAspAsnLysLysArgThrAlaLeu1leLysAlaValGlnCys 217
|||||
202 TGCTGTGACAAACGAAACAGACAGCTGTGATGAAGCTGTACAAATGCC 251
|||||
217 InGluAspGluCysAlaLeuMetLeuLeuGluH1sGlyThrAspProAsn 233
|||||
252 AGGAAAGGAATGTGCACTATCTGTAGAACACATGGCGATCCAAAT 301
|||||
234 IleProAspGluTyrGlyAsnThrLeuH1sTyrAla1leTyrAsnGly 250
|||||
302 CTTCGGATGTCCATGCAACACTGCTCTTCACATGCTGTCTATATGA 351
|||||
250 uAspLysLeuMetAlaLysAlaLeuLeuLeuTyrGlyValAsp1leGlu 267
|||||
352 GGAATATTCAGTACAAACAGCTGTTTGTATGATGACCAATATTTAG 401
|||||
267 erTyrAsnLysH1sGlyLeuThrProLeuLeuGlyValH1sGluGln 283
|||||
402 CAATAAACAAGATGATGACACACACTTTACTTCTGAGTAAAGAAA 451
|||||
284 LysGlnGlnValValLysPheLeu1leLysLysLysAlaAsnLeuAsnAl 300
|||||
452 AAGCAGCAATGTGTGAATTTTAAATAAGAAAAAGCAATGTAAATGTC 501
|||||
300 aleuAspArgTyrGlyArgThrAlaLeu1leleuAlaValCysCysGly 317
|||||
502 AGTAGATAAAG.....TTGGAA..... 511
|||||
317 erAlaSer1leValSerLeuLeuGluGlnAsn1leAspValSerSer 333
|||||
512 .....TTGGAA..... 517
|||||
334 GlnAspLeuSerGlyGlnThrAlaArgGluTyrAlaValSerSerH1s 350
|||||
518 .....ACCACTCAC.. 526

```

```

350 shIsval11leCysGlnleuSerAspTyrIysGlnIysGlnmetLeu 367
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
527 .....CAACTAATTTCAGAAATTAAGAGAAAGAGTACTTA 563
367 ysIseserSerGluAnSerAsn.....ProGluAn 377
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
564 AACATTCTTCTCAAAATAGTAATTCAGTGAGTAAGAGCTGGAAGACTTC 613
378 ValserArThrArGAsnIys 384
:|||||:
614 TTAAGACAGCTTCTTGCAAA 634

seq_name: gb_est1:AU120666

seq_documentation_block:
LOCUS AU120666 HEMBI Homo sapiens cDNA clone HEMBI001175 5', mRNA
DEFINITION AU120666 HEMBI Homo sapiens cDNA clone HEMBI001175 5', mRNA
sequence.
ACCESSION AU120666 GI:10935901
VERSION AU120666.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 843)
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,U., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1533-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3851
Fax: 81-438-52-3852
Email: genomese@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source 1..843
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEMBI001175"
/tissue="whole embryo, mainly body"
/dev_stage="embryo, 10 weeks"
/note="Vector: pME18SFL3"
BASE COUNT 240 a 178 c 245 g 176 t 4 others
ORIGIN
Alignment_scores:
Quality: 587.50 Length: 211
Ratio: 3.539 Gaps: 1
Percent Similarity: 78.673 Percent Identity: 56.398

alignment_block:
US-09-810-936-304 x AU120666 ..
Align seg 1/1 to: AU120666 from: 1 to: 843

90 AlameIysThrIeuArGAsnIysMetGlyIysTyrCysCysHisCysph 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
164 GCCATGAAGAAATTTTACAGAGGCGAGCGCCCTGGGCTCTT 213
106 eProCysCysArGIseryIysSerIysValGlyAlaTrrpGlyAspT 123
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
214 CGCGCGCGCGCG.....AGGAGCAGCGCGGAGCGGCGGCGAGC 254

```

```

123 yAspAspSerAlaPheMetGluProArGTYrHisValArgGlyGluAsp 139
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
255 CGGAGGAGGCGGCTACGACGCCGGGCTACACAGCTCCGAGACCGAGAT 304
140 LeuAspIysLeuHisArgAlaIaIaTrrpGlyIysValProArgIysAs 156
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
305 CTGGCAGATCCCAAAAGCTGCGAGCGGGGTAAATGGCGAAAGTGCA 354
156 PleuIleValMetLeuArGAspThrAspValAsnIysLysAspIysGln 173
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
355 GCAGATCTCTTCTGCTCAGCAAGAAATGGCTGACGATGACCAAGATGA 404
173 ySarGThrAlaLeuHisLeuAlaSerAlaAsnGlyAsnSerGluValAl 189
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
405 ACAGAGAGCGCTACATTTGGCGCTGGCCATGGTCATCCGAAAGTAGTA 454
190 LysLeuLeuLeuAspArGArGcysGlnIeuAsnValLeuAspAsnIys 206
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
455 ACTCTCCGTGGTGACAGAAATGCCAGCTCAATGCTGTGCAACGAAAA 504
206 sarGThrAlaLeuIleYsAlaValGlnCysGlnGluAspIuCyAlaL 223
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
505 CAGGACAGCGCTGATGAGAGCGTGACATCCAGAGAAAGAAATGTGCA 554
223 eumetLeuLeuGlnHisGlyThrAspProAsnIleProAspGluTyrGly 239
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
555 CTATTCTGCTAGAACATGGTGCTGATCCAAATCTTGGCGGATGTCCATGGC 604
240 AsnThrThrLeuHisTyrAlaIleTyrAsnGluAspIysLeuMetAla 256
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
605 AACACGCTCTTCATATGCTGCTATATATGAGGACATATCAGTAGCAAC 654
256 salaleuLeuLeuTyrGlyAlaAspIleGluSerIysAsnIysHisIyl 273
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
655 AAAGCTGCTTTGTATGATGCAAAATATGAAACAAAGACAAATGTGTGAA 704
273 eutrhProLeuLeuLeuGlyValHisGlnGlnIysGlnIvalValIys 289
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
705 TMAACACCACTTTTACTTGCAGTAGTGGAAGAAAGACAAATGTGTGAA 754
290 PheuleuIleLysIysLysAlaAsnLeuAsnAla 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
755 TTTTATATTAAGAAAAAACCAATGTAAATGCC 787

seq_name: gb_est2:BI871077

seq_documentation_block:
LOCUS BI871077 754 bp mRNA linear EST 11-OCT-2001
DEFINITION 603394250F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5403996 5',
mRNA sequence.
ACCESSION BI871077
VERSION BI871077.1 GI:16044750
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 754)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
Clone sequencing by: Incyte Genomics, Inc.
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
plate: LMNL2030 row: d column: 13
High quality sequence stop: 729.

```

```

FEATURES
    source
        Location/Qualifiers
            1..754
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:540396"
                /clone_lib="NIH_MGC_90"
                /issue_type="adenocarcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: Liver; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."
BASE COUNT      234 a      153 c      171 g      196 t
ORIGIN
alignment_scores:
    Quality: 538.00      Length: 248
    Ratio: 2.924      Gaps: 5
    Percent Similarity: 74.194      Percent Identity: 46.774
alignment_block:
    US-09-810-936-304 x B1871077 ..
Align seg 1/1 to: B1871077 from: 1 to: 754
133 TyHisValArgGlyIleuAspLeuHisArgAlaIleTPtr 149
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10 TATCATCTGAAGGGG.....ATCCACAGAGCTGTCTCTA 44
45 T.....CGTGAATCGAGAGACTGAAGTTCCTCTCTCA 79
149 pGlyIysValProArgLysAspLeu.....IleValMetLeu 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
162 rGAspThrAspValAlaAsnLysLysAspLysGlnLysArgThrAlaLeuHis 178
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 CGGCTTATGCATCATTAAGACAGACAGAGAAAGACCGCCCTGATCAT 129
179 IeuAlaSerLysAsnGlyAsnSerGluValAlaLysLeuLeuAspArg 195
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130 TTGGCCCTGCTCCACTGCGCCACCGGAAAGTGCATCTCTGCTGCTCAG 179
195 pArgCysGlnLeuAsnValLeuAspAsnLysLysArgThrAlaLeuIle 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
180 AAGTATGGAGCTTAACCTCTCGCAGCGTGAACACACACCTCTGATCA 229
212 YsaIaValAlaGlnCysGlnIleuAspGluCysAlaLeuMetLeuLeuHis 228
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
230 AGCGTGTACAACTGAGGACGAGAGCGTTGTGCACAACCTCTCTGCAAAAT 279
229 GLyThrAspProAsnIleProAspGlyIuArgLysAsnThrIleuHisTy 245
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
280 GGGCCCATCAATATATACGAGTGTCTTGGAGAGCTGCTGCTGCACTA 329
245 rAlaIleTyAsnGlnLysAspLysLeuMetAlaLysAlaLeuLeuTyrg 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
330 CGGTGTATATGAGATGATCATCATGATGAGAAAACCTCTTCATATG 379
262 IyAlaAspIleGlySerLysAsnLysHisGlyLeuThrProLeuLeuLeu 278
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
380 GTGCAAAATATTGAAGATGCGGAGAGTGAATATCCGCACTGTCTCTT 429
279 GLyValHisGlnGlnLysGlnGlnValValLysPheLeuIleTyLysLy 295
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
430 GCTGTAGTCAGAAAGAAAGTGAATGCGAATTTTATTATTAAGAAAAA 479
295 sAlaIAsnLeuAsnAlaLeuAspArgTyrgLysArgThrAlaLeuLeuA 312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
480 AGCAAAATATTAATGCGGTGATATTTGTGGCAGATCAGCCCTCATACATG 529
312 IaValCysGlySerAlaSerIleValSerLeuLeuLeuGlnGlnAsn 328
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

530 CTGTACTCTTGAGAAAAAGATATAGTCATCTTCTTCGACACAAAT 579
329 ILAspValSerSerGlnAspLeuSerGlyGlnThrAlaArg...GluTy 344
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
580 ATTAGATGTGTTTCTCGACCATGTGTAGACAAAGCTTGCAAGACGATTA 629
344 rAlaValSerSerHisHisValIleCysGlnLeuLeuSerAspTyrl 361
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
630 TGCCAGCGAGGCTTAAGACAGAGTCTTTTGAACATATCATATGACATATG 679
361 ysgIuIysGlnMetLeuLysIleSerSerGluAsnSerAsnPro 375
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
680 ACAGAAAGAAACATGAAGAGCTTCTATTA...AATAGCAATCCA 720
seq_name: gp_htc:AK017783
seq_documentation_block:
LOCUS      AK017783      1758 bp      mRNA      linear      HTC 19-JAN-2002
DEFINITION Mus musculus 8 days embryo whole body cDNA, RIKEN full-length
              enriched library, clone:5730521P14:homolog to to KIAA1074 PROTEIN,
              full insert sequence.
ACCESSION  AK017783
VERSION    AK017783.1 GI:12857204
KEYWORDS   HTC; CAP trapper.
SOURCE     Mus musculus (strain:O57Bt/6J) 8 days embryo cDNA to mRNA,
              clone_lib:RIKEN full-length enriched mouse cDNA library
              clone:5730521P14.
ORGANISM   Mus musculus
REFERENCE  1 (sites)
AUTHORS   Carninci,P., and Hayashizaki,Y.
TITLE     High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED    10349636
2 (sites)
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
              Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
              prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20493974
PUBMED    11042159
3 (sites)
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
              Kono,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
              Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishibe,T., Harada,A.,
              Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
              Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,M.,
              Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
              Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE     RIKEN integrated sequence analysis (RISA) system-384-format
              sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PUBMED    11076861
4 (sites)
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
              FANTOM Consortium.
TITLE     Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409, 685-690 (2001)
REFERENCE  5 (bases 1 to 1758)
AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
              Arakawa,T., Balderelli,R., Bono,H., Brownstein,M., Butt,C.,
              Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
              Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Hori,F.,
              Hume,D., Imotani,K., Ishii,Y., Itoh,M., Iwama,M., Kasukawa,T.,
              Kato,H., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Koya,S.,
              Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
              Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
              Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
              Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,

```

## TITLE

Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
Tanaka, T., Tejima, T., Toya, T., Yamamura, T., Yamana, K.,  
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and  
Hayashizaki, Y.

## Direct Submission

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp,  
URL: http://genome-gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

COMMENT  
Please visit our web site (http://genome-gsc.riken.go.jp/) for  
further details.

## FEATURES

## source

Location/Qualifiers  
1. 1758  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="MGI:1903054"  
/db\_xref="taxon:10090"  
/clone="5730521P14"  
/clone\_1lb="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="8 days embryo"  
159. .>1758  
/note="data source:SPRR, source key:Q9UP88, evidence:ISS  
homolog to KIAA1074 PROTEIN  
putative"  
/codon\_start=-1  
/protein\_id="BAB30930.1"  
/db\_xref="GI:12857205"  
/translation="MKKIFGRSKGSPPLGPSPARPRNCVFGESASGSHVPRYH  
DKDGRHKAAVGDVAKVOHILILKSGVNRDKDRTALHLCAVGHPEVTVLVE  
RKCDIARDESSSTALIKAVOCOECECALILDHDAEDVMSGNALHYAVYSEMT  
SMAKILAHANANIEAKNKDILPMLIAVENKQHTVEFLVKKASITAHVQDLSRGM  
FEVGRKRLQSSNSNPVNDSESGSTRSPYSPGADSPMPSEDEYFADKNPKTN  
LTERPTAAQOSRKNTKCGSEELDNARDDSDSESDAIEYEPASVYQVTLSS  
KQSPDYEGATELAIEBENGIDYIESASQEDPNNDLFRADGMRKSKMSALUL  
GEDEHSPWDSISVSLKVDVSGTAQDTKRRAHQILEVYTFPCMSGSS  
FKMAKLESRNVLPAVHMGMENKAPRYAIMEPTIERRAVLNKTETVGMTDAQTF  
KSEPSVSRREOTRLSGEDSDQQRV"

## CDS

## BASE COUNT

531 a 392 c 474 g 361 t

## ORIGIN

alignment\_scores:  
Quality: 521.50 Length: 386  
Ratio: 2.595 Gaps: 10  
Percent similarity: 52.073 Percent identity: 34.456

## alignment\_block:

us-09-810-936-304 x AK017783 ..

Align seg 1/1 to: AK017783 from: 1 to: 1758

```

3 ValGluValAspSerMetProAlaIaSerSerValLysLysProPheG1 19
||| ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| |||
132 GTGCGCTGTGAGGGCGCTACCC.....TTCGCATGAGAAGATTTTGG 175
19 yLeuArSerLysMetGlyLysTrpCysArgGlySerPheProCysCysA 36
|:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
176 CTTGAGGAGTTAAG..... 188
36 rgGluSerGlyLysSerAsnValGlyThrSerGlyAspHisAspAspSer 52
||| ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 .....GCCCATCGCCCTTAGCGCCCTTCGCC..... 215
53 AlawetLysThrLeuArSerLysMetGlyLysTrpCysArgHisCysPhe 69
||| ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| |||
216 .....CGCCCGCGCAGCAAC..... 230
69 eProCysCysArgGlySerGlyLysSerAsnValGlyLysSerGlyAspH 86
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
231 .....TGCCTGGCTTTGAGCGCAGAGT.....GCTAGTGGCTGCC 267

```

```

86 tsAspAspSerAlaMetLysThrLeuArgAsnLysMetGlyLysTrpCys 102
|||
268 AC..... 269
103 CysHisCysPheProCysCysArgGlySerGlyLysSerLysValGlyAl 119
269 ..... 269
119 aTrpGlyAspTrpAspAspSerAlaPheMetGluProArgTrpHisValA 136
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
270 .....GTGCCCATGATCCATATTC 288
136 rgGlyLysLysPheLysLysLeuHisArgAlaAlaTrpTrpGlyLysVal 152
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
289 ACACAGAGATATGGGCAGATCCACAGAGCTGCACAGCTGGAGCATGTA 338
153 ProArgLysAspLeuIleValMetLeuArgAspThrAspValAsnLysLys 169
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
339 GCCAAAGTCCAGATATCCCTATTCCTTGAAAGAGTGGCGAAGCATAG 388
169 sAspLysGlnLysArgThrAlaLeuHisLeuAlaSerAlaAsnGlyAsnS 186
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
389 AGACAGAGAGAGACAGAGACTGCTACATCTTGCTGCTGCTACGCGCAC 438
186 eArgLysValLysLeuLeuAspArgArgCysGlnLeuAsnValLeu 202
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
439 CAGAGAGTGTACTCTTATGTAGAGAGAAATGTGAGATAGATGCCCGT 488
203 AspAsnLysLysArgThrAlaLeuIleLysAlaValAlaGlnCysGlnLys 219
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
489 GACAGCGAGAGCAGCAGCGCCTCATTAAGCCCTGACAGTCCAGAGAGA 538
219 pGluCysAlaLeuMetLeuLeuGlnHisGlyThrAspProAsnLysProA 236
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
539 AGATGTGCGAGCATCTGCTGACACAGGCGCGCATCCAAATGTGATGG 588
236 spGluTrpGlyAsnThrThrLeuHisLysTrpAlaIleTrpAsnGlnLys 252
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
589 ACAGCAGCGGCGACACCGCCCTCCACTACGCGGCTACAGTGAAGAACG 638
253 LeuMetAlaLysAlaLeuLeuLeuTrpGlyAlaAspLysLysSerLys 269
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
639 TCCATGCGACCCAAACTGCTCCGACACCAATGCAAGTGAAGCAGAAA 688
269 pLysHisGlyLeuThrProLeuLeuLeuGlyValHisGlnLysGlnG 286
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
689 CAAAGATGACCTATACACCTATGTTACTGTAAAGGAAATGAACACAC 738
286 lnaValLysPheLeuIleLysLysLysAlaAsnLeuAsnAlaLeuAsp 302
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
739 ATATGTGGAATTTTATGTAAGAAAGAAACCAAGATATGACAGCATGAT 788
303 ArgTrpGlyArgThrAlaLeuIleLeuAlaValCysCysGlySerAla 319
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
789 CAGCTTGGAA..... 797
319 rIleValSerLeuLeuLeuGlnGlnAsnIleAspValSerSerGlnAsp 336
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
797 ..... 797
336 euSerGlyGlnThrAlaArgGlyTrpAlaValSerSerHisHisVal 352
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
798 .ACCAACAGCGCAAAATGTTGAAATAT..... 821
353 lIeCysGlnLeuLeuSerAspTrpLysGlnLysGlnMetLeuLysIle 369
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
822 .....GATGGAAGAGACTACAA.....AG 841
369 rSerGlnAsnSerAsnPro.....GluAsnValSerArgT 381
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
842 ATCTGAAGAAACGAAATCCAGTGAATAGTGTGAAGATGATGCTTTAA 891

```

381 hrrargan 383  
|||||:::  
892 CCAGCTCT 899

seq\_name: gb\_htc:AK015948

seq\_documentation\_block:  
LOCUS AK015948 1054 bp MRNA linear HTC 19-JAN-2002  
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493053J220;homolog to WGCSC:H.DJ1035002.1 PROTEIN (FRAGMENT), full insert sequence.  
ACCESSION AK015948  
VERSION AK015948.1 GI:12854490  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone:493053J220.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognath; Muridae; Murinae; Mus.  
REFERENCE Carninci,P. and Hayashizaki,Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE Carninci,P., Shibata,Y., Hayatsu,N., Sugihara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
REFERENCE Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsumi,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakauechi,S., Ikegami,T., Kashiwagi,K., Fujisake,Y., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,M., Toneyake,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.  
TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multipicillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
REFERENCE 4 (sites)  
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)  
REFERENCE 5 (bases 1 to 1054)  
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraka,T., Hoti,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurikara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numata,R., Ohtsuka,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,Y., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (10-Jul-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suenho-cho, Tsukuba, Ibaraki, 305-8572, Japan (E-mail:genome-test@gscc.riken.go.jp, kanagawa 230-0045, Japan)

URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216

COMMENT  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

**FEATURES**

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGGACAGCAGCATCCAGAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGCAGAGATTCGCAGCTAAATTAAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH109.

**Location/Qualifiers**

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="MGI:190851"
/db_xref="taxon:10090"
/clone="495032L20"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
110_949
/note="data source:SPR, source key:Q9UDM3, evidence:ISS
homolog to W05SC:H.DJ1035002.1 PROTEIN (FRAGMENT)"

```

/codon\_start=1  
 /protein\_id="BAB30047.1"  
 /db\_xref="GI:12854491"  
 /translation="MKKFPFPGKRRKRTDSSHSSSVPIISLAKTAPNLSIGGGYRLR  
 DKHLKHKRAATIGNEOKLDYLERKKYVANGDKSRPPLACANQYTNISLIE  
 NCKKINVDSENRPLIKAVECOESCATYLLHGADPMLVYVYNFLAHTAVCGNII  
 SIANKLLDYRKANLEAKMKGDHPPLLAVAPNNNNMKYFLIKKGADVYNSDKNHTTALM  
 IALIVPTSSVKLLIQDDTLAKHDIDYGFAEYASFNQFTWHTITANNENKKRTQD  
 TAY"  
 polyA\_signal 1033..1038  
 /note="putative"  
 polyA\_site 1054  
 /note="putative"  
 BASE COUNT 346 a 209 c 207 g 290 t 2 others  
 ORIGIN

[illegible]

```
alignment_block:
US-09-810-936-304 x AK015948 ..
```

Align seg 1/1 to: AK015948 from: 1 to: 1054

72 CysArggLySerGlyLySSerAsnValGlyAlaSerGlyAspHisAspAs 88  
|||||::: ||::: :: ||::: ||

88 pSerAlaMetLysThrLeuArgAsn.....LysMetG 99

64 GCTGTTGGTTAGGATTC CAGGAACATCTTACTT AGAGATCTTACCATGA 113

114 AGAGTTC.....TTCGCC..TTCAGGGGAAAAAGGAGACA 148



```

116 LysValGlyAlaTrpGlyAspTyrAspSer..... 126
149 .....GATGATTCACACAGCACTCCTCAGA 174
127 .....AlaPheMetLupro..... 131
175 AGTCCCATTAAGCTGGCCAAACCCACCGCCCAACCTGTCCATAGGTG 224
132 ..ATGTyrHisValArgGlyGluAspLeuAspLysLeuHisArgAlaIa 147
225 GAGATATTCGCTTCGAGATAGCACTTAAGAACTTCATTAACCTCCT 274
148 TrpTrpGlyValProArgLysAspLeuIleValMetLeuArgAspTh 164
275 ACCATTGGCAATAGCAGCAAGCTGAAGATTAACCTTGAACGCAAAATA 324
164 rAspValAsnLysLysAspLysGlnLysArgThrAlaLeuHisLeuAla 181
325 CAATGTGAATGGGGGACAAAGAGACAGAACACCTTGCTTGCGCT 374
181 erAlaAsnGlyAsnSerGluValValLysLeuLeuLeuAspArgArgCys 197
375 GTCGTATGGATATACAAATATGTCTCTCTTATAGATCAATATGC 424
198 GlnLeuAsnValLeuAspAsnLysLysArgThrAlaLeuIleLysAla 214
425 AAATTTATGTCCAGTACTGAAACAGACCCATTCATTAAGCAGT 474
214 LglCysGlnGluAspGluCysAlaLeuMetLeuLeuGluHisGlyThra 231
475 AGAGTGTCAACAGGAGTGTGTACGGTCTCTCTTCCTTCATGGGTGAG 524
231 sPrProAsnIleProAspGluTyrGlyAsnThrThrLeuHisTyrAlaIle 247
525 ATCCAATTCGTGTAGATGTTATAGTAATACGGCTCCATATAGCTGTT 574
248 TyrAsnGluAspLysLeuMetAlaLysAlaLeuLeuTyrGlyAlaAs 264
575 TGTGGCCCAAAATATTCATTAAGCAATTAATGCTTCATTAAGCCAA 624
264 pIleGluSerLysAsnLysHisGlyLeuThrProLeuLeuGluGlyAla 281
625 TCTTGAAGCCAAATATAGAGCGTCACTCGCTTTACTTGCTGTG 674
281 lsgLuglnLysGlnGlnValValLysPheLeuIleLysLysAlaAsn 297
675 CTGAAATTAATGAAATATGTAAATTCCTCTGAGAAAGACACAGAT 724
298 LeuAsnAlaLeuAspArgTyrGlyArgThrAlaLeuIleLeuAlaValCy 314
725 GTAATGATCAGATTAATAACACAGACACCATCATGATGCTGTGAT 774
314 scyGlySerAlaSerIleValSerLeuLeuGluGlnAsnIleAspV 331
775 TGTGTAACCAACAAGTCTGTAAACTTCTCTCAAGCAAGATAGGAGC 824
331 alSerSerGlnAspLeuSerGlyGlnThrAlaArgGluTyrAla..... 345
825 TAGCCCAACAAGATATTATGATTACAGCTGAGGAATATGCTTCATTT 874
346 .....ValSerSerHisHisValIleCysGlnLeuLeuSerAspTy 360
875 AATGGCTTTACTATGTATACCAT..... 898
360 rLysGluLysGlnMetLeuLysIleSerSerGlnLysSerAsnProGlu 377
899 .....ATACCTGCAATATATGAAACAGAGAA 926
377 snValSerArgThr 381
927 AATCAGAGCAACA 940
seq_name: gb_est2:Bg720647

```

```

seq_documentation_block:
LOCUS      Bg720647              694 bp    mRNA    linear    EST 08-MAY-2001
DEFINITION 602692528P1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4824630 5',
            mRNA sequence.
ACCESSION  Bg720647
VERSION    Bg720647.1  GI:13999834
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 694)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgraps-remail.nih.gov
            Tissue Procurement: Miklos Palcovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
            Toshitaki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLNL0736 row: h column: 07
            High quality sequence stop: 694.
FEATURES
     source
         1..694
             location/Qualifiers
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:4824630"
                 /clone_lib="NIH_MGC_97"
                 /lab_host="DH10B"
                 /note="Organ: testis; Vector: pBluescript (modified
                 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                 ); Oligo-dT primed using primer 5'-ATTGTTTATTTTATVN-3',
                 size-selected for average insert size 2.2 kb and
                 normalized to 5x. This is a primary library enriched
                 for full-length clones and constructed using the
                 Cap-trapper method (Carninci, in preparation). Library
                 constructed by M. Brownstein (NIH/NHGRI, National
                 Institutes of Health). Note: this is a NIH-MGC library."
BASE COUNT      108 a      173 c      228 g      185 t
ORIGIN
alignment_scores:
    Quality: 496.00      Length: 105
    Ratio: 5.167        Gaps: 0
    Percent Similarity: 91.429      Percent Identity: 83.810
alignment_block:
US-09-810-936-304 x Bg720647 ..
Align seg 1/1 to: Bg720647 from: 1 to: 694
1 MetValValGluValAspSerMetProAlaAlaSerSerValLysPyr 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
376 ATGGTGGCTGAGGTTGTTCATATGCCGCTGCTGTGCTGAACAAGCC 425
426 ATTGATCTCAGAGCAAGATGGCAAGTGTGCTCCACACCGCTGCCCT 475
17 oPheGlyLeuArgSerLysMetGlyLysTrpCysCysArgCysPheProC 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
426 ATTGATCTCAGAGCAAGATGGCAAGTGTGCTCCACACCGCTGCCCT 475
34 yscysArgGluSerGlyLysSerAsnValGlyThrSerGlyLysPheLysp 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
476 GCTGAGGGGGAGCGCAAGACACACTGTGTGAGACACACGAC 525
51 AspSerAlaMetLysThrLeuArgSerLysMetGlyLysTrpCysArgH 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
526 GACCTCTTATGAGACGCTCAGGAGCAAGATGGCAAGTGTGCCACCA 575

```

















726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bernaldo. "

BASE COUNT 156 a 137 c 138 g 108 t

ORIGIN

alignment\_scores:

Quality: 468.00 Length: 160

Ratio: 3.345 Gaps: 0

Percent Similarity: 82.500 Percent Identity: 58.125

alignment\_block:

US-09-810-936-304 x AA910780 ..

Align seg 1/1 to: AA910780 from: 1 to: 539

```

121 GUAAPRTYTRAPSPSERALAPHEMETGLUPROAGTYHISVALARGCL 137
||||| : : : : : : : : : : : : : : : : : : : : : : :
43 GCGAGACCGGGGAGGCGCTACTCGACCGCGCTACACGCTCGAGA 92
||||| : : : : : : : : : : : : : : : : : : : : : : :
137 YGLUASPLEUASPLYLEUHHISATGALAATATPTTPGLYLSVALPROA 154
:::||||| : : : : : : : : : : : : : : : : : : : : : :
93 CGAGATCTCGGCAAGATCCACAAAGCTGCCAGCGCGGTAAATGTGGCA 142
||||| : : : : : : : : : : : : : : : : : : : : : : :
154 IGLYASPLEUILEVALMETLEUARGSPRHRSPVALASNLYSLSASP 170
::: : : : : : : : : : : : : : : : : : : : : : :
143 AAGTGACGACGATCTTTGCTCAGGAAGATGGCTTGAACGATAGAGAC 192
||||| : : : : : : : : : : : : : : : : : : : : : : :
171 LYSGLINDYASGTTHRALALEUHHISLEUALSERALASNLGYASNSERG 187
||||| : : : : : : : : : : : : : : : : : : : : : : :
193 AAGATGACAGACGCGCTCTACATTTGGCCTG. TGCATGTGCATGCAGA 241
||||| : : : : : : : : : : : : : : : : : : : : : : :
187 UVALVALYLSLEULEUASPARARGCYSGINLEUASNLVALLEUASPA 204
||||| : : : : : : : : : : : : : : : : : : : : : : :
242 AGTAGTACTCTCGTGTGACAGAAATGCGACCTCAATGTGTGGACA 291
||||| : : : : : : : : : : : : : : : : : : : : : : :
204 SNLYLSYARGTHRALALEUILEYLSALAVAGINCYSGLINLUASPLU 220
||||| : : : : : : : : : : : : : : : : : : : : : : :
292 ACGAAACAGGACACCTGTGATGAGAGCCTGACAAATGCCAGAAAGAAA 341
||||| : : : : : : : : : : : : : : : : : : : : : : :
221 CYSALALEUETLEULEUHHISGLYTHASPRASNLLEPRASPL 237
||||| : : : : : : : : : : : : : : : : : : : : : : :
342 TGTGCACATTTCTGCTAGAACATGTCGATGCCAAATTTTCCGATGT 391
||||| : : : : : : : : : : : : : : : : : : : : : : :
237 ULYTGLYASNTHTHTLEUHHISTYRALALEYTRASNLSPYLSLEUM 254
:::||||| : : : : : : : : : : : : : : : : : : : : : :
392 CCATGGCAACACTGCTCTTCACTATGCTGTAAATGAGACATATACAG 441
||||| : : : : : : : : : : : : : : : : : : : : : : :
254 ETALALYSALALEUULEUYTGLYALASPLIEGUSELTSASNLYS 270
:::||||| : : : : : : : : : : : : : : : : : : : : : :
442 TAGCAACAAGCTGCTTTTGTATGATGCAAAATNTGAGACAAACCAAG 491
||||| : : : : : : : : : : : : : : : : : : : : : : :
271 HISGLYLEUTHPROLEULEUENGLIYVAL 280
||||| : : : : : : : : : : : : : : : : : : : : : : :
492 GATGACCTTCACACACCTTTTACTTGACAGTA 521
||||| : : : : : : : : : : : : : : : : : : : : : : :

```

seq\_name: gb\_est2:BG716974

seq\_documentation\_block:

LOCUS BG716974 916 bp mRNA linear EST 08-MAY-2001

DEFINITION 60268005F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4821576 5',

mrna sequence.

ACCESSION BG716974 GI:13996161

VERSION BG716974.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 916)

AUTHORS NIH-MGC http://mgi.mgi.nhl.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgabrs-r@mail.nih.gov  
Tissue Procurement: Miklos Palikovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM10728 row: 1 column: 01  
High quality sequence stop: 832.  
Location/Qualifiers

## FEATURES

## source

```

1..916
/organism="Homo sapiens"
/ld_xref="taxon:9606"
/clone="IMAGE:4821576"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
): Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."

```

## BASE COUNT

282 a 174 c 218 g 242 t

## ORIGIN

alignment\_scores:

Quality: 468.00 Length: 202

Ratio: 3.000 Gaps: 0

Percent Similarity: 77.228 Percent Identity: 46.040

alignment\_block:

US-09-810-936-304 x BG716974 ..

Align seg 1/1 to: BG716974 from: 1 to: 916

```

130 GUPTROATGTYTHHISVALARGLYGLUASPLEUASPLYLEUHHISARGAL 146
:::||||| : : : : : : : : : : : : : : : : : : : : : :
161 CAGCCAGGCTACCAACCTTCGAGAAAGATTTAAAGAACTTCACAGAGC 210
||||| : : : : : : : : : : : : : : : : : : : : : : :
146 AATATPTTPCLYLSVALPROARGLYSASPLEUILEVALMETLEUARGA 163
||||| : : : : : : : : : : : : : : : : : : : : : : :
211 TCTCTCAGTCCGGGATTTGAAGAGCTGAAGGAATACCTTCAGATCAAGA 260
||||| : : : : : : : : : : : : : : : : : : : : : : :
163 SPTRHASPVALASNLYSASPLYSGLINDYARGTHRALALEUHHISLEU 179
||||| : : : : : : : : : : : : : : : : : : : : : : :
261 AATATGATGTAAATATGCGAGCAAAAATATCGAAGACCTTTGCACTCA 310
||||| : : : : : : : : : : : : : : : : : : : : : : :
180 AATASERALAASNLGYASNSERGLUVALVALYLSLEULEUASPARARG 196
||||| : : : : : : : : : : : : : : : : : : : : : : :
311 GCCGTGCTAATGACATACAGATGTTGACTCTTCTCTAATTGAGACACA 360
||||| : : : : : : : : : : : : : : : : : : : : : : :
196 GYSGSLINLEUASNLVALLEUASPAENLYSARGTHRALALEUILEYSA 213
:::||||| : : : : : : : : : : : : : : : : : : : : : :
361 ATGCAAAATTAATGTCGGGATAGTGAACAAATCCCATTTGATTAAGG 410
||||| : : : : : : : : : : : : : : : : : : : : : : :
213 LAVALGINCYSGLINLUASPLIUCYSALALEUETLEUENGLIHSGLY 229
||||| : : : : : : : : : : : : : : : : : : : : : : :
411 CAGTACAGTGCACAAATGAGGATGTCTCTATCTTCTTAAACCTTGGT 460
||||| : : : : : : : : : : : : : : : : : : : : : : :
230 TTRASPPROASNLLEPROASPLIUTYGLYASNTHTHTLEUHHISTYRAL 246
:::||||| : : : : : : : : : : : : : : : : : : : : : :
461 GCGAGCCAGATCTGAGGATATTCGTTAATATATGTCCTTCACTATGCG 510
||||| : : : : : : : : : : : : : : : : : : : : : : :
246 ALLEYTASNLGYASPLYSLEUETALALYSALALEUETLEUYTGLY 263
||||| : : : : : : : : : : : : : : : : : : : : : : :
511 TCTTTGTGTCACAAAGTTTGTCATTAAGTTGAAGAAACGCTTGAATACGAG 560
||||| : : : : : : : : : : : : : : : : : : : : : : :

```

